b UNIVERSITÄT BERN

hp-DGFEM for second-order mixed elliptic problems in polyhedra

D. Schötzau, C. Schwab, and T. P. Wihler

Research Report 2013-10

7.11.2013

Mathematics Institute
University of Bern
Sidlerstrasse 5
CH-3012 Bern
Switzerland

www.math.unibe.ch

hp-DGFEM FOR SECOND-ORDER MIXED ELLIPTIC PROBLEMS IN POLYHEDRA

DOMINIK SCHÖTZAU, CHRISTOPH SCHWAB, AND THOMAS P. WIHLER

Abstract. We prove exponential rates of convergence of hp-dG interior penalty (IP) methods for second-order elliptic problems with mixed boundary conditions in polyhedra which are based on axiparallel, σ-geometric anisotropic meshes of mapped hexahedra and anisotropic polynomial degree distributions of μ-bounded variation. Compared to homogeneous Dirichlet boundary conditions in [10, 11], for problems with mixed Dirichlet-Neumann boundary conditions, we establish exponential convergence for a nonconforming dG interpolant consisting of elementwise L^2 projections onto elemental polynomial spaces with possibly anisotropic polynomial degrees, and for solutions which belong to a larger analytic class than the solutions considered in [11]. New arguments are introduced for exponential convergence of the dG consistency errors in elements abutting on Neumann edges due to the appearance of non-homogeneous, weighted norms in the analytic regularity at corners and edges. The nonhomogeneous norms entail a reformulation of dG flux terms near Neumann edges, and modification of the stability and quasi-optimality proofs, and the definition of the anisotropic interpolation operators. The exponential convergence results for the piecewise L^2 projection generalizes [10, 11] also in the Dirichlet case.

1. Introduction

Consider an open, bounded polyhedron $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\Gamma = \partial\Omega$ that consists of a finite union of plane faces $\Gamma_{\boldsymbol{\iota}}$ indexed by $\boldsymbol{\iota} \in \mathcal{J}$. The sets $\Gamma_{\boldsymbol{\iota}}$ are assumed to be bounded, plane polygons whose sides form the (open) edges of Ω. The set $\{\Gamma_{\boldsymbol{\iota}}\}_{\boldsymbol{\iota}\in J}$ is partitioned into two sets $\mathcal{J}_D$ and $\mathcal{J}_N$ of Dirichlet and of Neumann faces, respectively, i.e., $\mathcal{J} = \mathcal{J}_D \,\dot\cup\, \mathcal{J}_N$, with disjoint union. Then we consider the diffusion equation

$$-\Delta u = f \qquad \text{in } \Omega, \tag{1.1}$$

$$\gamma_0(u) = 0 \qquad \text{on } \Gamma_{\boldsymbol{\iota}} \subset \partial\Omega, \quad \boldsymbol{\iota} \in \mathcal{J}_D, \tag{1.2}$$

$$\gamma_1(u) = 0 \qquad \text{on } \Gamma_{\boldsymbol{\iota}} \subset \partial\Omega, \quad \boldsymbol{\iota} \in \mathcal{J}_N, \tag{1.3}$$

where the operators γ_0 and γ_1 denote the trace and (co)normal derivative operators, respectively. With the Sobolev space $V := H^1_D(\Omega) := \{v \in H^1(\Omega) : v|_{\Gamma_{\boldsymbol{\iota}}} = 0,\ \boldsymbol{\iota} \in \mathcal{J}_D\}$ and the continuous bilinear form $a(u,v) := \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}\boldsymbol{x}$, the variational form of problem (1.1)–(1.3) is to find $u \in H^1_D(\Omega)$ such that

$$a(u,v) = \int_\Omega f v \, \mathrm{d}\boldsymbol{x} \qquad \forall v \in H^1_D(\Omega)\ . \tag{1.4}$$

For every $f \in V^\star = H^1_D(\Omega)^\star$, the dual space of V, problem (1.4) admits a weak solution $u \in H^1_D(\Omega)$. The solution is unique if $\mathcal{J}_D \neq \emptyset$, and unique up to constants if $\mathcal{J}_D = \emptyset$ (in which case we also require the compatibility condition $\langle f, 1\rangle_{V\times V^\star} = 0$).

This paper is a continuation of our work [10, 11] on hp-version discontinuous Galerkin (dG) finite element methods (FEM) for second-order elliptic boundary-value problems in polyhedral domains $\Omega \subset \mathbb{R}^3$. In [10], we showed the well-posedness, stability and quasi-optimality of hp-version interior penalty (IP) discontinuous Galerkin discretizations of (1.1) in the pure Dirichlet case when $\mathcal{J} = \mathcal{J}_D$, $\mathcal{J}_N = \emptyset$, and the homogeneous essential boundary conditions (1.2) are posed

---

2010 *Mathematics Subject Classification.* 65N30.

Key words and phrases. hp-dGFEM, second-order elliptic problems in 3D polyhedra, mixed Dirichlet-Neumann boundary conditions, exponential convergence.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), the European Research Council AdG grant STAHDPDE 247277, and the Swiss National Science Foundation (SNF).

on all of $\partial\Omega$. In [11], we then used these results to prove exponential rates of convergence in the number of degrees of freedom, on appropriate combinations of σ-geometric meshes and $\mathfrak{s}$-linearly increasing anisotropic elemental polynomial degrees; see also [14] for related work on linear elasticity.

In this work, we consider the case $\mathcal{J}_N \neq \emptyset$. The case $\mathcal{J}_N = \emptyset$ is the pure Dirichlet case where exponential convergence was established in [10, 11]. The *hp*-error analysis in the present paper is along the lines of [10, 11], however, there are some significant differences: as shown in [3], the solutions of mixed Dirichlet-Neumann or pure Neumann problems for second order, elliptic boundary value problems in polyhedral domains with piecewise analytic data belong to analytic classes specified in terms of countably normed Sobolev spaces. In elements in the vicinity of $\Gamma_{\boldsymbol{\iota}}$, for $\boldsymbol{\iota} \in \mathcal{J}_D$, the analytic classes coincide with those for the Dirichlet case, and accordingly, exponential convergence would follow as in [11]. In the present paper, we provide an alternative proof also in the Dirichlet case, constructing an *hp*-interpolant from *elementwise L^2-projections*. The exponential convergence proofs in this work will focus on stability and exponential convergence bounds in elements in the vicinity of $\Gamma_{\boldsymbol{\iota}}, \boldsymbol{\iota} \in \mathcal{J}_N$. Here, new technical difficulties (as compared to [11]) arise, due to the solutions belonging to countably normed Sobolev spaces with *nonhomogeneous weights* $N^m_{\boldsymbol{\beta}}(\Omega)$ introduced in [3]. In the case of homogeneous Dirichlet conditions (i.e., when $\mathcal{J}_N = \emptyset$), these spaces coincide with the (smaller) spaces $M^m_{\boldsymbol{\beta}}(\Omega)$ for which we proved exponential convergence in [10, 11]. When $\mathcal{J}_N \neq \emptyset$, however, we have the *strict inclusion* $N^m_{\boldsymbol{\beta}}(\Omega) \supsetneq M^m_{\boldsymbol{\beta}}(\Omega)$, due to the different structure of the weights near Neumann edges, i.e., edges at the intersection of two faces $\Gamma_{\boldsymbol{\iota}}$, $\boldsymbol{\iota} \in \mathcal{J}_N$. Compared to [10, 11], the different structure of the weights entails essential modifications in the definition of the anisotropic *hp*-interpolation operators, as well as in the error bounds in elements containing Neumann faces. Compared to the results of [10, 11], we present here an *hp*-dG discretization for (1.1)–(1.3) with $\mathcal{J}_N \neq \emptyset$. On axiparallel hexahedral meshes and for linear anisotropic polynomial degree distributions, and for isotropic diffusion coefficients, we establish exponential convergence. Specifically, we show that the *hp*-dG approximations are well-defined, satisfy the Galerkin orthogonality property and, hence, the dG energy error can be bounded with respect to a suitable discontinuous elemental polynomial interpolation operator. We generalize the result in [11] (for the case $\mathcal{J}_N = \emptyset$), and prove that *hp*-dGFEM achieve *exponential convergence*, i.e., asymptotic convergence rate bounds of the form $C\exp(-b\sqrt[5]{N})$, where N is the number of degrees of freedom, and where $b, C > 0$ are independent of N.

The outline of the article is as follows: In Section 2, we recapitulate regularity results in countably normed weighted Sobolev spaces for the solution of (1.1) – (1.3) from [3], extending the pioneering work [2] in two dimensions to the three-dimensional case. In Section 3, we define *hp*-dG finite element spaces on *σ-geometric meshes* of mapped hexahedral elements with possibly anisotropic *polynomial degree distributions*. In Section 4, we extend the stability and quasi-optimality results of [10] to the mixed boundary conditions considered here. Particular attention is being paid to the analysis of consistency errors in elements abutting at “Neumann-edges”, being edges where two faces with homogeneous Neumann boundary conditions meet. In Section 5, we present exponential convergence bounds for the consistency terms arising in the dG-stability analysis, and state our exponential convergence result (Theorem 5.6). Sections 6–7.6 are devoted to the proof of this result. Although we use ideas and notation from [10, 11], the proof of exponential convergence in the present paper is self-contained, and the results in several respects stronger than the analysis in [11]: exponential convergence is shown for larger classes of solutions, and for an (quasi)interpolant which requires merely L^2-regularity of the solution, thereby generalizing the analysis in [11]. This is purchased at the expense of additional powers of the maximal polynomial degree (as compared to [11]) appearing in the consistency error bounds; these are subsequently absorbed into the exponentially small terms.

The notation employed throughout this paper is consistent with [10, 11]. In particular, we shall frequently use the function

$$\Psi_{q,r} = \frac{\Gamma(q+1-r)}{\Gamma(q+1+r)}, \qquad 0 \leq r \leq q,\ q, r \in \mathbb{N}, \tag{1.5}$$

where Γ is the Gamma function satisfying $\Gamma(m+1) = m!$, for any $m \in \mathbb{N}$. Moreover, we shall use the notations "$\lesssim$" or "$\simeq$" to mean an inequality or an equivalence containing generic positive multiplicative constants which are independent of the local mesh sizes, polynomial degrees, and regularity parameters, as well as of the geometric refinement level, but which may depend on the geometric refinement ratio σ and on the linear polynomial degree slope $\mathfrak{s}$.

2. Regularity

To establish exponential convergence of *hp*-dG methods, it is necessary to specify the precise regularity of solutions of (1.1)–(1.3) in countably normed weighted Sobolev spaces. To do so, we follow [3], based on the notations already introduced in [10, 11].

2.1. **Subdomains and Weights.** We denote by $\mathcal{C}$ the set of corners $\boldsymbol{c}$, and by $\mathcal{E}$ the set of open edges $\boldsymbol{e}$ of Ω. The singular set of Ω is then given by

$$\mathcal{S} = \left(\bigcup_{\boldsymbol{c} \in \mathcal{C}} \boldsymbol{c} \right) \cup \left(\bigcup_{\boldsymbol{e} \in \mathcal{E}} \boldsymbol{e} \right) \subset \Gamma. \tag{2.1}$$

For $\boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{e} \in \mathcal{E}$, and $\boldsymbol{x} \in \Omega$, we define the following distance functions:

$$r_{\boldsymbol{c}}(\boldsymbol{x}) = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{c}), \qquad r_{\boldsymbol{e}}(\boldsymbol{x}) = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{e}), \qquad \rho_{\boldsymbol{ce}}(\boldsymbol{x}) = r_{\boldsymbol{e}}(\boldsymbol{x}) / r_{\boldsymbol{c}}(\boldsymbol{x}). \tag{2.2}$$

We *assume* the vertices of Ω to be separated:

$$\exists\, \varepsilon(\Omega) > 0 : \qquad \bigcap_{\boldsymbol{c} \in \mathcal{C}} B_\varepsilon(\boldsymbol{c}) = \emptyset, \tag{2.3}$$

where $B_\varepsilon(\boldsymbol{c})$ denotes the open ball in $\mathbb{R}^3$ with center $\boldsymbol{c}$ and radius ε. For each corner $\boldsymbol{c} \in \mathcal{C}$, $\mathcal{E}_{\boldsymbol{c}} = \{ \boldsymbol{e} \in \mathcal{E} : \boldsymbol{c} \cap \overline{\boldsymbol{e}} \neq \emptyset \}$ denotes the set of all edges of Ω which meet at $\boldsymbol{c}$. Similarly, for any $\boldsymbol{e} \in \mathcal{E}$, the set of corners of $\boldsymbol{e}$ is given by $\mathcal{C}_{\boldsymbol{e}} \equiv \partial \boldsymbol{e} = \{ \boldsymbol{c} \in \mathcal{C} : \boldsymbol{c} \cap \overline{\boldsymbol{e}} \neq \emptyset \}$. Then, for $\boldsymbol{c} \in \mathcal{C}$, $\boldsymbol{e} \in \mathcal{E}$ and $\boldsymbol{e}_{\boldsymbol{c}} \in \mathcal{E}_{\boldsymbol{c}}$, we define the neighborhoods

$$\begin{aligned} \omega_{\boldsymbol{c}} &= \{ \boldsymbol{x} \in \Omega : r_{\boldsymbol{c}}(\boldsymbol{x}) < \varepsilon \ \wedge \ \rho_{\boldsymbol{ce}}(\boldsymbol{x}) > \varepsilon \quad \forall \boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}} \}, \\ \omega_{\boldsymbol{e}} &= \{ \boldsymbol{x} \in \Omega : \ r_{\boldsymbol{e}}(\boldsymbol{x}) < \varepsilon \ \wedge \ r_{\boldsymbol{c}}(\boldsymbol{x}) > \varepsilon \quad \forall \boldsymbol{c} \in \mathcal{C}_{\boldsymbol{e}} \}, \\ \omega_{\boldsymbol{ce}_{\boldsymbol{c}}} &= \{ \boldsymbol{x} \in \Omega : r_{\boldsymbol{c}}(\boldsymbol{x}) < \varepsilon \ \wedge \ \rho_{\boldsymbol{ce}_{\boldsymbol{c}}}(\boldsymbol{x}) < \varepsilon \}. \end{aligned} \tag{2.4}$$

Possibly by reducing ε in (2.3), we may partition the domain Ω into four *disjoint* parts,

$$\overline{\Omega} = \overline{\Omega_0 \,\dot\cup\, \Omega_{\mathcal{C}} \,\dot\cup\, \Omega_{\mathcal{E}} \,\dot\cup\, \Omega_{\mathcal{CE}}}, \tag{2.5}$$

where

$$\Omega_{\mathcal{C}} = \bigcup_{\boldsymbol{c} \in \mathcal{C}} \omega_{\boldsymbol{c}}, \qquad \Omega_{\mathcal{E}} = \bigcup_{\boldsymbol{e} \in \mathcal{E}} \omega_{\boldsymbol{e}}, \qquad \Omega_{\mathcal{CE}} = \bigcup_{\boldsymbol{c} \in \mathcal{C}} \bigcup_{\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}} \omega_{\boldsymbol{ce}}. \tag{2.6}$$

We shall refer to the subdomains $\Omega_{\mathcal{C}}$, $\Omega_{\mathcal{E}}$ and $\Omega_{\mathcal{CE}}$ as *corner*, *edge* and *corner-edge neighborhoods* of Ω, respectively, and the remaining *interior part* of the domain Ω is defined by $\Omega_0 := \Omega \setminus \overline{\Omega_{\mathcal{C}} \cup \Omega_{\mathcal{E}} \cup \Omega_{\mathcal{CE}}}$.

In the sequel, it will be useful to refine the partition in (2.6) by introducing the following subsets of $\mathcal{C}$ and $\mathcal{E}$, respectively:

$$\begin{aligned} \mathcal{C}_D &:= \left\{ \boldsymbol{c} \in \mathcal{C} \, : \, \exists \mathfrak{s} \in \mathcal{J}_D \text{ with } \boldsymbol{c} \cap \overline{\Gamma}_{\mathfrak{s}} \neq \emptyset \right\}, \\ \mathcal{E}_D &:= \left\{ \boldsymbol{e} \in \mathcal{E} \, : \, \exists \mathfrak{s} \in \mathcal{J}_D \text{ with } \boldsymbol{e} \cap \overline{\Gamma}_{\mathfrak{s}} \neq \emptyset \right\}, \\ \mathcal{E}_N &:= \mathcal{E} \setminus \mathcal{E}_D \, . \end{aligned} \tag{2.7}$$

Corners in $\mathcal{C}_D$ and edges in $\mathcal{E}_D$ abut at at least one Dirichlet face $\Gamma_{\boldsymbol{\iota}}$ for $\boldsymbol{\iota} \in \mathcal{J}_D$. Note that we possibly have $\mathcal{E}_N = \emptyset$. Hence, the edge neighborhoods in (2.6) can be further partitioned into:

$$\Omega_{\mathcal{E}} = \Omega_{\mathcal{E}_D} \,\dot\cup\, \Omega_{\mathcal{E}_N}, \tag{2.8}$$

where, as in (2.6), we let $\Omega_{\mathcal{E}_D} = \bigcup_{\boldsymbol{e} \in \mathcal{E}_D} \omega_{\boldsymbol{e}}$, and $\Omega_{\mathcal{E}_N} = \bigcup_{\boldsymbol{e} \in \mathcal{E}_N} \omega_{\boldsymbol{e}}$.

2.2. **Weighted Sobolev Spaces.** To each $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{E}$ we associate a corner and an edge exponent $\beta_{\boldsymbol{c}}, \beta_{\boldsymbol{e}} \in \mathbb{R}$, respectively. We collect these quantities in the multi-exponent

$$\boldsymbol{\beta} = \{\beta_{\boldsymbol{c}} : \boldsymbol{c} \in \mathcal{C}\} \cup \{\beta_{\boldsymbol{e}} : \boldsymbol{e} \in \mathcal{E}\} \in \mathbb{R}^{|\mathcal{C}|+|\mathcal{E}|}. \tag{2.9}$$

Inequalities of the form $\boldsymbol{\beta} < 1$ and expressions like $\boldsymbol{\beta} \pm s$, where $s \in \mathbb{R}$, are to be understood componentwise. For example, $\boldsymbol{\beta} + s = \{\beta_{\boldsymbol{c}} + s : \boldsymbol{c} \in \mathcal{C}\} \cup \{\beta_{\boldsymbol{e}} + s : \boldsymbol{e} \in \mathcal{E}\}$. We shall often use the notation

$$b_{\boldsymbol{c}} = -1 - \beta_c, \quad \boldsymbol{c} \in \mathcal{C}\ , \qquad b_{\boldsymbol{e}} = -1 - \beta_{\boldsymbol{e}} \quad \boldsymbol{e} \in \mathcal{E}\ . \tag{2.10}$$

At the heart of the exponential convergence analysis of hp-approximations in three dimensions is the analytic regularity of the solution u of (1.1)–(1.2) near the set of edges $\mathcal{E}$ of Ω. In order to describe it, we recall from [10], for corners $\boldsymbol{c} \in \mathcal{C}$ and edges $\boldsymbol{e} \in \mathcal{E}$, the local coordinate systems in $\omega_{\boldsymbol{e}}$ and $\omega_{\boldsymbol{ce}}$ which are chosen such that $\boldsymbol{e}$ corresponds to the direction $(0,0,1)$. Then, we denote quantities that are transversal to $\boldsymbol{e}$ by $(\cdot)^\perp$, and quantities parallel to $\boldsymbol{e}$ by $(\cdot)^\|$. In particular, if $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^\perp, \alpha^\|)$ with $\boldsymbol{\alpha}^\perp = (\alpha_1, \alpha_2)$ and $\alpha^\| = \alpha_3$ is a multi-index corresponding to the three local coordinate directions in a subdomain $\omega_{\boldsymbol{e}}$ or $\omega_{\boldsymbol{ce}}$, then the operator $\mathsf{D}^{\boldsymbol{\alpha}}$ denotes the partial derivative in these local coordinate directions. Likewise notation shall be employed below in anisotropic quantities related to a face. Furthermore, we will write $|\boldsymbol{\alpha}^\perp| = \alpha_1 + \alpha_2$, and $|\boldsymbol{\alpha}| = |\boldsymbol{\alpha}^\perp| + \alpha_3$

The solution u of (1.1)–(1.3) belongs to a scale $N^m_{\boldsymbol{\beta}}(\Omega)$ of countably normed spaces which are, in the case $\mathcal{J}_N \neq \emptyset$ under consideration here, strictly larger than the scale $M^m_{\boldsymbol{\beta}}(\Omega)$ of spaces considered in [10, 11] for the pure Dirichlet case, i.e., for $\mathcal{J} = \mathcal{J}_D$, so that the exponential convergence results proved in this paper generalize those in [10, 11]. We define the semi-norm

$$\begin{aligned}
&|u|^2_{N^k_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} := \\
&\sum_{\substack{\boldsymbol{\alpha}\in\mathbb{N}^3_0\\|\boldsymbol{\alpha}|=k}} \Bigg\{ \|\mathsf{D}^{\boldsymbol{\alpha}} u\|^2_{L^2(\Omega_0)} + \sum_{\boldsymbol{e}\in\mathcal{E}_D} \Big\| r_{\boldsymbol{e}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{e}})} + \sum_{\boldsymbol{e}\in\mathcal{E}_N} \Big\| r_{\boldsymbol{e}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{e}})} \\
&\quad + \sum_{\boldsymbol{c}\in\mathcal{C}_D} \Bigg(\Big\| r_{\boldsymbol{c}}^{\beta_c+|\boldsymbol{\alpha}|} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{c}})} + \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_{\mathrm{D}}} \Big\| r_{\boldsymbol{c}}^{\beta_c+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{ce}})} \\
&\qquad + \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_N} \Big\| r_{\boldsymbol{c}}^{\beta_c+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{ce}})} \Bigg) \\
&\quad + \sum_{\boldsymbol{c}\in\mathcal{C}\setminus\mathcal{C}_D} \Bigg(\Big\| r_{\boldsymbol{c}}^{\max\{\beta_c+|\boldsymbol{\alpha}|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{c}})} + \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_D} \Big\| r_{\boldsymbol{c}}^{\max\{\beta_c+|\boldsymbol{\alpha}|,0\}} \rho_{\boldsymbol{ce}}^{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{ce}})} \\
&\qquad + \sum_{\boldsymbol{e}\in\mathcal{E}_{\boldsymbol{c}}\cap\mathcal{E}_N} \Big\| r_{\boldsymbol{c}}^{\max\{\beta_c+|\boldsymbol{\alpha}|,0\}} \rho_{\boldsymbol{ce}}^{\max\{\beta_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u\Big\|^2_{L^2(\omega_{\boldsymbol{ce}})} \Bigg) \Bigg\}.
\end{aligned} \tag{2.11}$$

For $m > k_{\boldsymbol{\beta}}$, with

$$k_{\boldsymbol{\beta}} := -\min\{\min_{\boldsymbol{c}\in\mathcal{C}} \beta_{\boldsymbol{c}}, \min_{\boldsymbol{e}\in\mathcal{E}} \beta_{\boldsymbol{e}}\}, \tag{2.12}$$

we denote by $N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)$ the space of functions u such that $\|u\|_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} < \infty$, with the norm $\|u\|^2_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)} := \sum_{k=0}^m |u|^2_{N^k_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)}$.

It follows from the definition of the norm $\|u\|_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)}$ that the spaces $N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)$ are monotonic with respect to the sets $\mathcal{C}_D$, $\mathcal{E}_D$: for $\emptyset \subseteq \mathcal{C}_D \subseteq \mathcal{C}$ and $\emptyset \subseteq \mathcal{E}_D \subseteq \mathcal{E}$, we have

$$M^m_{\boldsymbol{\beta}}(\Omega) := N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E}) \subseteq N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D) \subseteq N^m_{\boldsymbol{\beta}}(\Omega;\emptyset,\emptyset) =: N^m_{\boldsymbol{\beta}}(\Omega), \tag{2.13}$$

where $M^m_{\boldsymbol{\beta}}(\Omega)$ is the weighted Sobolev space obtained as the closure of $C^\infty_0(\Omega)$ with respect to the norm $\|\circ\|_{M^m_{\boldsymbol{\beta}}(\Omega)} = \|\circ\|_{N^m_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E})}$. For subdomains $K \subseteq \Omega$ we shall denote by $|\circ|_{N^k_{\boldsymbol{\beta}}(K;\mathcal{C}_D,\mathcal{E}_D)}$ the semi-norm (2.11) with all domains of integration replaced by their intersections with $K \subset \Omega$, and likewise we shall use the norm $\|\circ\|_{N^m_{\boldsymbol{\beta}}(K;\mathcal{C}_D,\mathcal{E}_D)}$.

2.3. **Analytic regularity of variational solutions.** We adopt the following classes of analytic functions from [3].

Definition 2.1. For subdomains $K \subseteq \Omega$ and any subsets $\mathcal{C}' \subset \mathcal{C}$, $\mathcal{E}' \subset \mathcal{E}$, the space $B_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')$ consists of all functions u such that $u \in N^m_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')$ for $m > k_{\boldsymbol{\beta}}$, with $k_{\boldsymbol{\beta}}$ as in (2.12), and such that there exists a constant $C_u > 0$ with the property that

$$|u|_{N^k_{\boldsymbol{\beta}}(K;\mathcal{C}',\mathcal{E}')} \leq C_u^{k+1} k! \qquad \forall k > k_{\boldsymbol{\beta}} \ . \tag{2.14}$$

Remark 2.2. The analytic class $B_{\boldsymbol{\beta}}(\Omega) = B_{\boldsymbol{\beta}}(\Omega;\emptyset,\emptyset)$ is closely related to the countably normed spaces $B^\ell_\beta(\Omega)$ introduced by Babuška and Guo in [2, 7, 8]: if the edge and corner exponents $\beta_{ij} \in (0,1)$ and $\beta_m \in (0, 1/2)$ introduced in [2, 7, 8] satisfy $\beta_{ij} = \beta_{\boldsymbol{e}} + \ell$ and $\beta_m = \beta_{\boldsymbol{c}} + \ell$ for every $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{E}$, then $B^\ell_\beta(\Omega) = B_{\boldsymbol{\beta}}(\Omega)$. By (2.13), we also have $A_{\boldsymbol{\beta}}(\Omega) = B_{\boldsymbol{\beta}}(\Omega;\mathcal{C},\mathcal{E})$, where $A_{\boldsymbol{\beta}}(\Omega)$ is the analytic class considered in [11].

We have the following regularity result (see [3, Theorem 7.3]).

Proposition 2.3. *There are bounds $b_{\mathcal{E}}, b_{\mathcal{C}} > 0$ (depending on Ω and on the space V) such that, for $\boldsymbol{b}$ satisfying*

$$0 < b_{\boldsymbol{c}} < b_{\mathcal{C}}, \quad 0 < b_{\boldsymbol{e}} < b_{\mathcal{E}}, \qquad \boldsymbol{e} \in \mathcal{E},\ \boldsymbol{c} \in \mathcal{C}, \tag{2.15}$$

any weak solution $u \in V$ defined in (1.4) *of problem* (1.1)–(1.3) *satisfies:*

$$f \in B_{1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D) \implies u \in B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D) \ . \tag{2.16}$$

Remark 2.4. We may and will assume in the following without loss of generality that in (2.15) there holds $0 < b_{\mathcal{C}},\ b_{\mathcal{E}} < 1$. Then $\beta_{\boldsymbol{c}}, \beta_{\boldsymbol{e}} \in (-2,-1)$ in (2.10). Consequently, we have $\kappa_\beta \in (1,2)$ in (2.12), and (2.14) holds for all $k > 1$. Moreover, for $|\boldsymbol{\alpha}^\perp| \geq 2$, there holds, $\max\{\beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|, 0\} = \beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|$.

Remark 2.5. Under the assumptions of Proposition 2.3, there holds

$$B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D) \subset C^0(\overline{\Omega}) \ . \tag{2.17}$$

This inclusion is a consequence of Remark 2.2 above on the equivalence of weighted analytic spaces defined via (2.11), (2.14), with the spaces of Babuška and Guo introduced in [7, 8], under our assumption that $0 < b_{\mathcal{C}}, b_{\mathcal{E}} \leq 1$ (cp. Remark 2.4); see [8, Theorem 5.10]. The assertions (2.16) and (2.17) imply in particular that point values of the solution $u \in B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ are well-defined at $\mathcal{E}$ and $\mathcal{C}$.

Remark 2.6. Note that the regularity (2.16) implies the a-priori estimates (2.14) in the weighted spaces with weights at *all* $\boldsymbol{c} \in \mathcal{C}$, even if $\boldsymbol{c}$ is a "Neumann corner", i.e. if only Neumann faces meet at corner $\boldsymbol{c}$. In the case of corners $\boldsymbol{c}$ of polyhedra in $\mathbb{R}^3$, corner weights do *not* imply homogeneous Dirichlet boundary conditions since by Hardy's inequality $\{u \in H^1(\Omega) :\ r_{\boldsymbol{c}}^{-1} u \in L^2(\Omega)\ \forall \boldsymbol{c} \in \mathcal{C}\} = H^1(\Omega)$ for bounded Lipschitz domains $\Omega \subset \mathbb{R}^3$. This implies that the Dirichlet corner weights do *not* contribute to the characterization of integrability of the weak solution $u \in V$ near the singular set $\mathcal{S}$ which is, by (2.11), completely characterized by the edge weight functions for all edges $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ which meet at corner $\boldsymbol{c} \in \mathcal{C}$. The regularity (2.16) in the analytic class $B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ implies $\mathcal{C}_D = \mathcal{C}$ in (2.11) so that only six out of the nine terms in the weighted semi-norms $|\cdot|_{N^k_{\boldsymbol{\beta}}(\Omega;\mathcal{C}_D,\mathcal{E}_D)}$ suffice to characterize the analytic regularity of u. In particular, the corner weights have the same structure as in the pure Dirichlet case, albeit with in general a larger range of the exponents $\beta_{\boldsymbol{c}}$, whereas for each edge $\boldsymbol{e} \in \mathcal{E}$, the two cases $\boldsymbol{e} \in \mathcal{E}_D$ and $\boldsymbol{e} \notin \mathcal{E}_D$ must be distinguished. The positivity of the indices $b_{\boldsymbol{e}}$, $b_{\boldsymbol{c}}$ in (2.15) implies with (2.10) that $-1 - \beta_{\mathcal{E}} < \beta_{\boldsymbol{e}} < -1$, $-1 - \beta_{\mathcal{C}} < \beta_{\boldsymbol{c}} < -1$, and $1 < k_{\boldsymbol{\beta}} < 1 + \min\{\beta_{\mathcal{C}}, \beta_{\mathcal{E}}\}$. Inspection of (2.11) reveals that this forces the solution to zero weakly *at Dirichlet edges* $\boldsymbol{e} \in \mathcal{E}_D$; however, the structure of the weights $r_{\boldsymbol{e}}^{\max\{\beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp|, 0\}}$ associated with Neumann edges $\boldsymbol{e} \notin \mathcal{E}_D$ in the third and sixth terms in (2.11) allows for nonzero traces of $u \in V$ at such edges.

3. hp-Subspaces in Ω

In [10], we constructed a class of hp-dG spaces on families $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell \geq 1}$ of *nested, σ-geometric meshes of hexahedral elements with ℓ layers of refinement*, polynomial degree distributions which are nonuniform, anisotropic within elements and $\mathfrak{s}$-linearly increasing between elements. Here, we recapitulate the construction in the particular case of *axiparallel* domains and meshes, and refer to [10, Section 3] for details and proofs.

3.1. **Geometric hp-Meshes in Ω.** We start from any coarse regular *quasiuniform* partition $\mathcal{M}^0 = \{Q_j\}_{j=1}^J$ of Ω into J convex axiparallel hexahedra. Each of these hexahedral elements $Q_j \in \mathcal{M}^0$ is the image under an affine mapping G_j of the *reference patch* $\widetilde{Q} = (-1,1)^3$, i.e. $Q_j = G_j(\widetilde{Q})$ for $j = 1, \ldots, J$. In fact, since the hexahedra $\{Q_j\}_j$ are assumed axiparallel, the mappings G_j are compositions of (isotropic) dilations and translations. Due to our assumption that the faces of Ω are plane, it is geometrically exact.

In [10], *canonical geometric mesh patches* on the reference patch $\widetilde{Q}$ have been constructed; see Figure 1. Geometric meshes towards corners and edges in Ω can then be obtained by again applying the patch mappings G_j to transform these canonical geometric mesh patches on the reference patch $\widetilde{Q}$ to the patches $Q_j \in \mathcal{M}^0$. It is important to note that the geometric refinements in the canonical patches have to be suitably selected and oriented in order to achieve a proper geometric refinement towards corners and edges of Ω. In addition, we allow for simultaneous geometric refinement towards several edges. In [10, Section 3.3], a specific construction of geometric meshes has been introduced in terms of four different hp-extensions (Ex1)–(Ex4) as displayed in Figure 1. They also apply to our exponential convergence analysis below. Moreover, the patches Q_j with $\overline{Q}_j \cap \mathcal{S} = \emptyset$ away from the singular support $\mathcal{S}$ are left unrefined, i.e., no refinement is considered on $\widetilde{Q}$.

Consider now the hexahedral patch $Q_j \in \mathcal{M}^0$. We denote the elements in the canonical geometric mesh patch associated with Q_j by $\widetilde{\mathcal{M}}_j = \{\widetilde{K}\}$, where we allow $\widetilde{\mathcal{M}}_j = \{\widetilde{Q}\}$ in the case of unrefined patches. The elements in $\widetilde{\mathcal{M}}_j$ are then transported to the physical domain Ω via the (finitely many) affine patch maps G_j. Moreover, for each $\widetilde{K} \in \widetilde{\mathcal{M}}_j$, we can write $\widetilde{K} = H_{j,\widetilde{K}}(\widehat{K})$, where $H_{j,\widetilde{K}} : \widehat{K} \to \widetilde{K}$ is a possibly anisotropic dilation combined with a translation of the reference cube $\widehat{K} = (-1,1)^3$ (to be distinguished from the reference patch $\widetilde{Q}$). Thus, the elements in the patch $Q_j \subset \Omega$ will be given by $\mathcal{M}_j = \left\{ K : K = (G_j \circ H_{j,\widetilde{K}})(\widehat{K}),\ \widetilde{K} \in \widetilde{\mathcal{M}}_j \right\}$, $j = 1, \ldots, J$. A geometric mesh in Ω is now given by

$$\mathcal{M} := \bigcup_{j=1}^{J} \mathcal{M}_j. \tag{3.1}$$

Throughout, we shall assume that the initial mesh $\mathcal{M}^0$ is sufficiently fine so that an element $K \in \mathcal{M}$ has non-trivial intersection with at most one corner $\boldsymbol{c} \in \mathcal{C}$ and at most one edge $\boldsymbol{e} \in \mathcal{E}$. By construction, each hexahedral element $K \in \mathcal{M}$ is the image of the reference cube $\widehat{K}$ under an element mapping Φ_K $K = \Phi_K(\widehat{K})$, which is a possibly anisotropic dilation with a translation from $\widehat{K}$ to K. We collect all element mappings Φ_K in the *mapping vector* $\boldsymbol{\Phi}(\mathcal{M}) := \{ \Phi_K : K \in \mathcal{M} \}$.

With each (axiparallel) element $K \in \mathcal{M}$ in the geometric mesh, let us associate a polynomial degree vector $\boldsymbol{p}_K = (p_{K,1}, p_{K,2}, p_{K,3}) \in \mathbb{N}_0^3$. Its components correspond to the coordinate directions in $\widehat{K} = \Phi_K^{-1}(K)$. The polynomial degree is called *isotropic* if $p_{K,1} = p_{K,2} = p_{K,3} = p_K$. We set $|\boldsymbol{p}_K| := \max_{i=1}^3 p_{K,i}$.

In the hp-error estimates, we shall often write K in the form

$$K := K^\perp \times K^\parallel, \tag{3.2}$$

where $K^\perp$ is an axiparallel rectangle of diameter $h_K^\perp$ in the first two coordinates $\boldsymbol{x}^\perp = (x_1, x_2)$ perpendicular to the nearest edge $\boldsymbol{e}$, and $K^\parallel$ is an interval of length $h_K^\parallel$ in the third coordinate direction $x^\parallel = x_3$ parallel to $\boldsymbol{e}$. Analogously, we then choose $p_{K,1} = p_{K,2} =: p_K^\perp$, and write $\boldsymbol{p}_K = (p_K^\perp, p_K^\parallel)$.

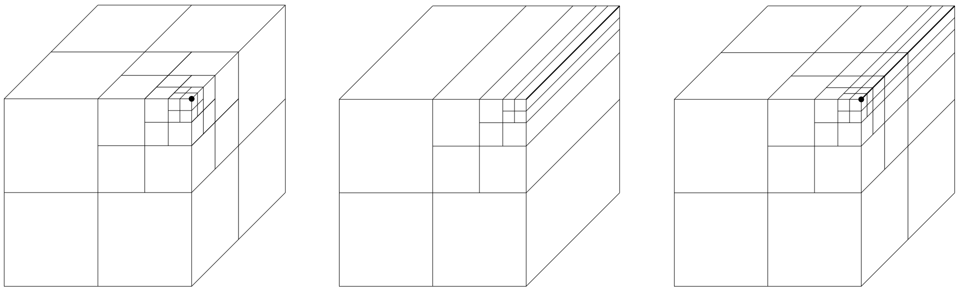

FIGURE 1. Examples of three basic geometric mesh subdivisions in the reference patch $\widetilde{Q}$ with subdivision ratio $\sigma = 1/2$: isotropic refinement towards the corner $\boldsymbol{c}$ (left), anisotropic refinement towards the edge $\boldsymbol{e}$ (center), and anisotropic refinement towards the edge-corner pair $\boldsymbol{ce}$ (right). The sets $\boldsymbol{c}, \boldsymbol{e}, \boldsymbol{ce}$ are shown in boldface.

Given a mesh $\mathcal{M}$ of hexahedral elements in Ω, we combine the elemental polynomial degrees $\boldsymbol{p}_K$ into the *polynomial degree vector* $\boldsymbol{p}(\mathcal{M}) := \{\, \boldsymbol{p}_K : K \in \mathcal{M} \,\}$, and define $\boldsymbol{p}_{\max} := \max_{K\in\mathcal{M}} |\boldsymbol{p}_K|$. We remark that, in addition to the mesh refinements, the extensions (Ex1)–(Ex4) introduced in [10] also provide appropriate polynomial degree distributions that increase $\mathfrak{s}$-linearly away from the singular set $\mathcal{S}$.

In the sequel, we shall be working with *sequences of σ-geometrically refined meshes* denoted by $\mathcal{M}_\sigma^{(0)}, \mathcal{M}_\sigma^{(1)}, \mathcal{M}_\sigma^{(2)}, \ldots, \mathcal{M}_\sigma^{(\ell)}, \ldots$, where $\mathcal{M}_\sigma^{(0)} := \mathcal{M}^0$. Here, $\sigma \in (0,1)$ is a fixed parameter defining the ratio of subdivision in the canonical geometric refinements in Figure 1. We shall refer to the index ℓ as *refinement level*, and to the sequence $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell\geq 1}$ as a *σ-geometric mesh family*; see [10, Definition 3.4].

3.2. **Mesh Layers.** As in [10, Section 3], we shall use the concept of *mesh layers:* these are partitions of $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell\geq 1}$ into certain subsets of elements with identical scaling properties in terms of their relative distance to the sets $\mathcal{C}$ and $\mathcal{E}$. The following result holds.

Proposition 3.1. *Any σ-geometric mesh family $\mathfrak{M}_\sigma$ obtained by iterating the basic hp-extensions (Ex1)–(Ex4) in* [10] *can be partitioned into a countable sequence of disjoint mesh layers $\{\mathfrak{L}_\sigma^j\}_{j=0}^{\ell-1}$, and a corresponding nested sequence of terminal layers $\mathfrak{T}_\sigma^\ell$, such that each $\mathcal{M}_\sigma^{(\ell)} \in \mathfrak{M}_\sigma$, $\ell \geq 1$, can be written as*

$$\mathcal{M}_\sigma^{(\ell)} = \mathfrak{L}_\sigma^0 \,\dot\cup\, \mathfrak{L}_\sigma^1 \,\dot\cup \ldots \dot\cup\, \mathfrak{L}_\sigma^{\ell-1} \,\dot\cup\, \mathfrak{T}_\sigma^\ell. \tag{3.3}$$

Elements in the submesh

$$\mathfrak{O}_\sigma^\ell := \mathfrak{L}_\sigma^0 \,\dot\cup\, \mathfrak{L}_\sigma^1 \,\dot\cup \ldots \dot\cup\, \mathfrak{L}_\sigma^{\ell-1} \subset \mathcal{M}_\sigma^{(\ell)} \in \mathfrak{M}_\sigma, \qquad \ell \geq 1, \tag{3.4}$$

are bounded away from $\mathcal{C} \cup \mathcal{E}$, while all elements in the terminal layer $\mathfrak{T}_\sigma^\ell$ have a nontrivial intersection with $\mathcal{C} \cup \mathcal{E}$. Evidently, $\mathcal{M}_\sigma^{(\ell)} = \mathfrak{O}_\sigma^\ell \,\dot\cup\, \mathfrak{T}_\sigma^\ell$ for $\ell \geq 1$.

We partition $\mathfrak{O}_\sigma^\ell$ into discrete corner, edge and corner-edge neighborhoods as $\mathfrak{O}_\sigma^\ell = \mathfrak{O}_\mathcal{C}^\ell \,\dot\cup\, \mathfrak{O}_\mathcal{E}^\ell \,\dot\cup\, \mathfrak{O}_{\mathcal{CE}}^\ell \,\dot\cup\, \mathfrak{O}_{\mathrm{int}}^\ell$, where for $\ell \geq 1$,

$$\begin{aligned}
\mathfrak{O}_{\mathrm{int}}^\ell &:= \left\{ K \in \mathfrak{O}_\sigma^\ell \,:\, \overline{K} \cap \Omega_0 \neq \emptyset \right\}, \\
\mathfrak{O}_\mathcal{C}^\ell &:= \left\{ K \in \mathfrak{O}_\sigma^\ell \,:\, \overline{K} \cap \Omega_\mathcal{C} \neq \emptyset \right\} \setminus \mathfrak{O}_{\mathrm{int}}^\ell, \\
\mathfrak{O}_\mathcal{E}^\ell &:= \left\{ K \in \mathfrak{O}_\sigma^\ell \,:\, \overline{K} \cap \Omega_\mathcal{E} \neq \emptyset \right\} \setminus (\mathfrak{O}_{\mathrm{int}}^\ell \cup \mathfrak{O}_\mathcal{C}^\ell), \\
\mathfrak{O}_{\mathcal{CE}}^\ell &:= \left\{ K \in \mathfrak{O}_\sigma^\ell \,:\, \overline{K} \cap \Omega_{\mathcal{CE}} \neq \emptyset \right\} \setminus \left(\mathfrak{O}_{\mathrm{int}}^\ell \cup \mathfrak{O}_\mathcal{C}^\ell \cup \mathfrak{O}_\mathcal{E}^\ell\right) .
\end{aligned} \tag{3.5}$$

Note that there exists $\ell_0 \geq 1$ (depending on ε from (2.3) and on σ) such that $\mathfrak{O}_{\mathrm{int}}^\ell = \mathfrak{O}_{\mathrm{int}}^{\ell_0}$ for $\ell \geq \ell_0$. Without loss of generality, we shall assume that the initial mesh is sufficiently fine so that we can choose $\ell_0 = 2$. Consequently, in what follows we shall simply write $\mathfrak{O}_{\mathrm{int}}$ instead of $\mathfrak{O}_{\mathrm{int}}^{\ell_0}$. In addition, we may assume without loss of generality that $\mathfrak{L}_\sigma^0 \subset \mathfrak{O}_{\mathrm{int}}^\ell$ for $\ell \geq \ell_0 = 2$.

For an element $K \in \mathcal{M}$, we set $h_K := \operatorname{diam}(K)$, and denote by $h_K^{\perp}$ and $h_K^{\|}$ the elemental diameters of K transversal respectively parallel to the singular edge $\boldsymbol{e} \in \mathcal{E}$ nearest to K; cp. [10]. For isotropic elements, we have $h_K^{\|} \simeq h_K^{\perp} \simeq h_K$. In a sequence $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell \geq 1}$ of σ-geometric meshes, we define for any $K \in \mathcal{M}_\sigma^{(\ell)}$, $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{E}$ the quantities:

$$d_K^{\boldsymbol{e}} := \operatorname{dist}(K, \boldsymbol{e}) = \inf_{\boldsymbol{x} \in K} r_{\boldsymbol{e}}(\boldsymbol{x}), \qquad d_K^{\boldsymbol{c}} := \operatorname{dist}(K, \boldsymbol{c}) = \inf_{\boldsymbol{x} \in K} r_{\boldsymbol{c}}(\boldsymbol{x}). \tag{3.6}$$

These quantities are closely related to the elemental diameters $h_K^{\perp}$ and $h_K^{\|}$; cp. [11, Prop. 3.2 & 3.4]. In particular, if $K = K^{\perp} \times K^{\|} \in \mathfrak{O}_\sigma^\ell$ as in (3.2), then $d_K^{\boldsymbol{c}} \simeq h_K^{\|}$, and $d_K^{\boldsymbol{e}} \simeq h_K^{\perp}$.

Similarly, we partition the terminal layer $\mathfrak{T}_\sigma^\ell$ into $\mathfrak{T}_\sigma^\ell := \mathfrak{T}_{\mathcal{C}}^\ell \,\dot\cup\, \mathfrak{T}_{\mathcal{E}}^\ell$, where

$$\mathfrak{T}_{\mathcal{C}}^\ell := \bigcup_{\boldsymbol{c} \in \mathcal{C}} \mathfrak{T}_{\boldsymbol{c}}^\ell, \qquad \mathfrak{T}_{\boldsymbol{c}}^\ell := \{\, K \in \mathfrak{T}_\sigma^\ell \,:\, \overline{K} \cap \boldsymbol{c} \neq \emptyset \,\}, \quad \boldsymbol{c} \in \mathcal{C}, \tag{3.7}$$

$$\mathfrak{T}_{\mathcal{E}}^\ell := \bigcup_{\boldsymbol{e} \in \mathcal{E}} \mathfrak{T}_{\boldsymbol{e}}^\ell, \qquad \mathfrak{T}_{\boldsymbol{e}}^\ell := \{\, K \in \mathfrak{T}_\sigma^\ell \setminus \mathfrak{T}_{\mathcal{C}}^\ell \,:\, (\overline{K} \cap \boldsymbol{e})^\circ \text{ is an entire edge of } K \,\}, \quad \boldsymbol{e} \in \mathcal{E}\,. \tag{3.8}$$

For $\mathcal{M}^0$ sufficiently fine, we may assume that $\mathfrak{T}_{\boldsymbol{c}}^\ell$ consists of at most a finite number (independent of $\boldsymbol{c} \in \mathcal{C}$, σ, and ℓ) of elements $K \in \mathfrak{T}_\sigma^\ell$. According to [11, Proposition 3.2], these corner elements $K \in \mathfrak{T}_{\boldsymbol{c}}^\ell$ are isotropic with $h_K \simeq h_K^{\perp} \simeq h_K^{\|} \simeq \sigma^\ell$, while elements in $K \in \mathfrak{T}_{\boldsymbol{e}}^\ell$ may be anisotropic with $d_K^{\boldsymbol{e}} \lesssim h_K^{\perp} \simeq \sigma^\ell$, and $d_K^{\boldsymbol{c}} \simeq h_K^{\|} \simeq \sigma^{\ell+1-j}$ for an exponent $2 \leq j \leq \ell + 1$.

3.3. **Finite Element Spaces.** Let $\mathcal{M} = \mathcal{M}_\sigma^{(\ell)}$, for some ℓ, be a geometric mesh of a σ-geometric mesh family $\mathfrak{M}_\sigma$ in Ω. Furthermore, let $\boldsymbol{\Phi}(\mathcal{M})$ and $\boldsymbol{p}(\mathcal{M})$ be the associated element mapping and elemental polynomial degree vectors, as introduced above. We then introduce the discontinuous *hp* finite element space

$$V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) = \left\{\, u \in L^2(\Omega) \,:\, u|_K \in \mathbb{Q}_{\boldsymbol{p}_K}(K),\, K \in \mathcal{M} \,\right\}. \tag{3.9}$$

Here, we define the local polynomial approximation space $\mathbb{Q}_{\boldsymbol{p}_K}(K)$ as follows: first, on the reference element $\widehat{K}$ and for a polynomial degree vector $\boldsymbol{p} = (p_1, p_2, p_3) \in \mathbb{N}_0^3$, we introduce the anisotropic polynomial space: $\mathbb{Q}^{\boldsymbol{p}}(\widehat{K}) = \mathbb{P}_{p_1}(\widehat{I}) \otimes \mathbb{P}_{p_2}(\widehat{I}) \otimes \mathbb{P}_{p_3}(\widehat{I}) = \operatorname{span}\{\, \hat{\boldsymbol{x}}^{\boldsymbol{\alpha}} \,:\, \alpha_i \leq p_i,\ 1 \leq i \leq 3 \,\}$. Here, for $p \in \mathbb{N}_0$, we denote by $\mathbb{P}^p(\widehat{I})$ the space of all polynomials of degree at most p on the reference interval $\widehat{I} = (-1, 1)$. Then, if K is a hexahedral element of $\mathcal{M}$ with associated elemental mapping $\Phi_K : \widehat{K} \to K$ and polynomial degree vector $\boldsymbol{p}_K = (p_{K,1}, p_{K,2}, p_{K,3})$, we define $\mathbb{Q}_{\boldsymbol{p}_K}(K) := \left\{ u \in L^2(K) : (u|_K \circ \Phi_K) \in \mathbb{Q}_{\boldsymbol{p}_K}(\widehat{K}) \right\}$. In the case where the polynomial degree vector $\boldsymbol{p}_K$ associated with K is isotropic, i.e., $p_{K,1} = p_{K,2} = p_{K,3} = p_K$, we simply write $\mathbb{Q}_{\boldsymbol{p}_K}(K) = \mathbb{Q}_{p_K}(K)$. For technical reasons that will become clear in the analysis, we will assume throughout the paper that all polynomial degrees on elements $K \in \mathfrak{O}_\sigma^\ell$ are greater than or equal to 3.

We now introduce two families of *hp*-finite element spaces for the discontinuous Galerkin methods; both yield exponentially convergent approximations and are based on the σ-geometric mesh families $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell \geq 1}$. The *first family of hp-dG subspaces* is defined by

$$V_\sigma^\ell := V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})), \qquad \ell \geq 1, \tag{3.10}$$

where the elemental polynomial degree vectors p_K in $\boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})$ are isotropic and uniform, given on each element $K \in M_\sigma^{(\ell)}$ as $p_K = \max\{3, \ell\}$. The *second family of hp-dG subspaces* is chosen as

$$V_{\sigma,\mathfrak{s}}^\ell := V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})), \qquad \ell \geq 1, \tag{3.11}$$

for an increment parameter $\mathfrak{s} > 0$. Here the polynomial degree vectors $\boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})$ are linearly increasing with slope $\mathfrak{s}$ away from $\mathcal{S}$, i.e., specifically, the polynomial degrees $\boldsymbol{p}_K^{\perp}$ and $p_K^{\|}$ within each element $K \in \mathcal{M}_\sigma^{(\ell)}$ increase linearly with the number of mesh layers between that element and the closest edge $\boldsymbol{e} \in \mathcal{E}$ respectively the closed corner $\boldsymbol{c} \in \mathcal{C}$ of Ω, with the factor of proportionality ("slope" in the terminology of [6]) being $\mathfrak{s} > 0$; see [10, Section 3]. In the pure Neumann case ($\mathcal{J}_D = \emptyset$) we consider the factor space $\widetilde{V}_{\sigma,\mathfrak{s}}^\ell = V_{\sigma,\mathfrak{s}}^\ell / \mathbb{R}$.

Remark 3.2. By construction, increasing the index j in the mesh layers $\mathfrak{L}^j_\sigma$ corresponds to moving from *inside the domain towards the singular set* $\mathcal{S}$, with $\mathfrak{L}^0_\sigma$ being the most inner layer, and the terminal layer $\mathfrak{T}^\ell_\sigma$ being the most outer layer abutting at $\mathcal{S}$; see (3.3). While this numbering takes into account the scaling properties of $\mathfrak{L}^j_\sigma$, it is in contrast to the notion of $\mathfrak{s}$-linearly increasing polynomial degrees where the polynomial degree increases $\mathfrak{s}$-linearly away from the *singular set into the interior of the domain*; see also [10].

3.4. **Properties of bounded variation.** The spaces V^ℓ_σ and $V^\ell_{\sigma,\mathfrak{s}}$ defined in (3.10) and (3.11), respectively, satisfy bounded variation properties with respect to the local mesh sizes and polynomial degrees. These properties will be implicitly used in our analysis. To describe them, let $\mathfrak{M}_\sigma$ be the underlying σ-geometric mesh family. For any $\mathcal{M} \in \mathfrak{M}_\sigma$, we define the set of all interior faces in $\mathcal{M}$ by

$$\mathcal{F}_I(\mathcal{M}) := \{\, f = (\partial K^\flat \cap \partial K^\sharp)^\circ \neq \emptyset \,:\, K^\flat, K^\sharp \in \mathcal{M} \,\}.$$

The set of all Dirichlet boundary faces is given by $\mathcal{F}_D(\mathcal{M}) := \{\, f = (\partial K \cap \partial \Gamma_{\boldsymbol{\iota}})^\circ \neq \emptyset \,:\, \boldsymbol{\iota} \in \mathcal{J}_D \,\}$, and similarly, we denote by $\mathcal{F}_N(\mathcal{M})$ the set of all Neumann faces. In addition, let $\mathcal{F}(\mathcal{M}) = \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}) \cup \mathcal{F}_N(\mathcal{M})$ denote the set of all (smallest) faces of $\mathcal{M}$. Furthermore, for an element $K \in \mathcal{M}$, we denote the set of its faces by $\mathcal{F}_K = \{\, f \in \mathcal{F}(\mathcal{M}) \,:\, f \subset \partial K \,\}$. For $K \in \mathcal{M}$ and $f \in \mathcal{F}_K$, we denote by $h^\perp_{K,f}$ the height of K over the face f, i.e., the diameter of element K in the direction transversal to f. Similarly, we denote by $p^\perp_{K,f}$ the polynomial degree of $\boldsymbol{p}_K$ transversal to f (defined as the corresponding component of $\Phi_K^{-1}(K)$).

The geometric mesh family now satisfies the following property with respect to the local mesh sizes: there is a constant $\mu_1 \in (0,1)$ only depending on σ and $\mathcal{M}^0$ such that

$$\mu_1 \,\leq\, h^\perp_{K^\sharp,f} / h^\perp_{K^\flat,f} \,\leq\, \mu_1^{-1}, \tag{3.12}$$

for all interior faces $f \in \mathcal{F}_I(\mathcal{M})$, and $\mathcal{M} \in \mathfrak{M}_\sigma$. Further, the family of degree vectors $\boldsymbol{p}_2(\mathcal{M}^{(\ell)}_\sigma)_{\ell \geq 1}$ introduced in (3.11) satisfies a similar property with respect to the polynomial degree: there is a constant $\mu_2 \in (0,1)$ (depending on $\mathfrak{s}$) such that $\mu_2 \,\leq\, p^\perp_{K^\sharp,f} / p^\perp_{K^\flat,f} \,\leq\, \mu_2^{-1}$, for all interior faces $f = \mathcal{F}_I(\mathcal{M}^{(\ell)}_\sigma)$, and $\ell \geq 1$. Note that for the family $\boldsymbol{p}_1(\mathcal{M}^{(\ell)}_\sigma)_{\ell \geq 1}$ in (3.10) this property is trivially satisfied.

4. Discontinuous Galerkin Discretization

In this section we present the hp-dG discretizations of (1.1)–(1.2) for which we shall prove exponential convergence. In addition, we shall adapt the stability and approximation results from [10, Section 4] to mixed boundary conditions. Throughout, $\mathcal{M} \in \mathfrak{M}_\sigma$ denotes a generic σ-geometric mesh.

4.1. **Trace operators and trace discretization parameters.** We shall first recall the jump and average operators over faces; cp. [10, 11]. For this purpose, consider an interior face $f \in \mathcal{F}_I(\mathcal{M})$ shared by two elements $K^\sharp, K^\flat \in \mathcal{M}$. Furthermore, let v respectively $\boldsymbol{w}$ be a scalar respectively vector-valued function that is sufficiently smooth inside the elements $K^\sharp, K^\flat$. Then we define the following jumps and averages of v and $\boldsymbol{w}$ along f:

$$\begin{aligned}
[\![v]\!] &= v|_{K^\sharp}\boldsymbol{n}_{K^\sharp} + v|_{K^\flat}\boldsymbol{n}_{K^\flat} & \langle\!\langle v \rangle\!\rangle &= {}^1\!/\!{}_2\,(v|_{K^\sharp} + v|_{K^\flat}) \\
[\![\boldsymbol{w}]\!] &= \boldsymbol{w}|_{K^\sharp} \cdot \boldsymbol{n}_{K^\sharp} + \boldsymbol{w}|_{K^\flat} \cdot \boldsymbol{n}_{K^\flat} & \langle\!\langle \boldsymbol{w} \rangle\!\rangle &= {}^1\!/\!{}_2\,(\boldsymbol{w}|_{K^\sharp} + \boldsymbol{w}|_{K^\flat})\,.
\end{aligned}$$

Here, for an element $K \in \mathcal{M}$, we denote by $\boldsymbol{n}_K$ the outward unit normal vector on ∂K. For a Dirichlet boundary face $f \in \mathcal{F}_D(\mathcal{M})$ belonging to $K \in \mathcal{M}$, we let $[\![v]\!] = v|_K \boldsymbol{n}_\Omega$, $[\![\boldsymbol{w}]\!] = \boldsymbol{w}|_K \cdot \boldsymbol{n}_\Omega$, and $\langle\!\langle v \rangle\!\rangle = v|_K$, $\langle\!\langle \boldsymbol{w} \rangle\!\rangle = \boldsymbol{w}|_K$, where $\boldsymbol{n}_\Omega$ is the outward unit normal vector on $\partial\Omega$.

Moreover, we define the trace discretization parameters $\mathtt{h}, \mathtt{p} \in L^\infty(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))$ by

$$\mathtt{h}_f := \mathtt{h}|_f := \min\big\{h^\perp_{K^\sharp,f}, h^\perp_{K^\flat,f}\big\}, \qquad \mathtt{p}_f := \mathtt{p}|_f := \max\big\{p^\perp_{K^\sharp,f}, p^\perp_{K^\flat,f}\big\}, \tag{4.1}$$

for any interior face $f \in \mathcal{F}_I(\mathcal{M})$ shared by $\partial K^\sharp$ and $\partial K^\flat$. For a Dirichlet boundary face $f \in \mathcal{F}_D(\mathcal{M})$ shared by ∂K and $\Gamma_{\boldsymbol{\iota}}$, $\boldsymbol{\iota} \in \mathcal{J}_D$, we set accordingly $\mathtt{h}_f := \mathtt{h}|_f = h^\perp_{K,f}$, $\mathtt{p}_f := \mathtt{p}|_f = p^\perp_{K,f}$.

4.2. hp**-IP dGFEM.** The problem (1.1)–(1.3) will be discretized using an interior penalty (IP) discontinuous Galerkin finite element method. Let $V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$ be an hp-dG finite element space on a σ-geometric mesh $\mathcal{M} \in \mathfrak{M}_\sigma$, with a degree vector $\boldsymbol{p}(\mathcal{M})$. For a fixed parameter $\theta \in \mathbb{R}$, we define the hp-discontinuous Galerkin solution u_{DG} by

$$u_{\mathrm{DG}} \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}): \qquad a_{\mathrm{DG}}(u_{\mathrm{DG}}, v) = \int_\Omega f v \,\mathrm{d}\boldsymbol{x} \qquad \forall\, v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}), \tag{4.2}$$

where the bilinear form $a_{\mathrm{DG}}(u, v)$ is given by

$$\begin{aligned} a_{\mathrm{DG}}(u, v) := \int_\Omega \nabla_h u \cdot \nabla_h v \;\mathrm{d}\boldsymbol{x} - \int_{\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \langle\!\langle \nabla_h w \rangle\!\rangle \cdot [\![v]\!] \,\mathrm{d}s \\ + \theta \int_{\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \langle\!\langle \nabla_h v \rangle\!\rangle \cdot [\![w]\!] \,\mathrm{d}s + \gamma \int_{\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \mathtt{j}\, [\![v]\!] \cdot [\![w]\!] \,\mathrm{d}s. \end{aligned}$$

Here, ∇_h is the elementwise gradient operator, and $\gamma > 0$ is a stabilization parameter that will be chosen sufficiently large. Furthermore, $\mathtt{j}$ is facewise defined as

$$\mathtt{j}\,|_f = \mathtt{p}_f^2 \mathtt{h}_f^{-1}, \qquad f \in \mathcal{F}_I(\mathcal{M}) \cap \mathcal{F}_D(\mathcal{M}). \tag{4.3}$$

Finally, the parameter θ allows us to describe a whole range of interior penalty methods: for $\theta = -1$ we obtain the standard symmetric interior penalty (SIP) method while for $\theta = 1$ the non-symmetric (NIP) version is obtained; cp. [1] and the references therein.

To address the well-posedness of the hp-dGFEM, we use the standard dG norm defined by

$$|\!|\!| v |\!|\!|_{\mathrm{DG}}^2 = \int_\Omega |\nabla_h v|^2 \;\mathrm{d}\boldsymbol{x} + \gamma \int_{\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \mathtt{j}\, |[\![v]\!]|^2 \;\mathrm{d}s, \tag{4.4}$$

for any $v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) + H^1(\Omega)$. In the pure Neumann case ($\mathcal{F}_D(\mathcal{M}) = \emptyset$), $|\!|\!| \cdot |\!|\!|_{\mathrm{DG}}$ is a norm on the subspace $(V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p}) + H^1(\Omega))/\mathbb{R}$.

4.3. **Galerkin orthogonality and stability properties.** In order to show the well-posedness of the dG formulation (4.2), we recall first the anisotropic trace inequality from [10, Lemma 4.2]:

Lemma 4.1. *Let $\mathcal{M} \in \mathfrak{M}_\sigma$, $0 < \sigma < 1$, $K \in \mathcal{M}$, $f \in \mathcal{F}_K$. For $1 \le q < \infty$ there exists $C_q > 0$ such that for any $v \in W^{1,q}(K)$ holds*

$$\|v\|_{L^q(f)}^q \le C_q \left(h_{K,f}^\perp\right)^{-1} \left(\|v\|_{L^q(K)}^q + \left(h_{K,f}^\perp\right)^q \|\partial_{K,f,\perp} v\|_{L^q(K)}^q \right) . \tag{4.5}$$

The constant $C_q > 0$ is independent of the element size and of the element aspect ratio, and $\partial_{K,f,\perp}$ signifies the partial derivative with respect to the (local coordinate) direction transversal to $f \in \mathcal{F}_K$.

Secondly, the following *Galerkin orthogonality* is crucial in the subsequent dG error analysis.

Proposition 4.2. *Suppose that the solution u of* (1.1)–(1.3) *belongs to $N^2_{-1-\boldsymbol{b}}(\Omega; \mathcal{C}, \mathcal{E}_D)$, where $\boldsymbol{b}$ is the weight vector from* (2.15)*. Then, the dG approximation $u_{\mathrm{DG}} \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$ from* (4.2) *satisfies $a_{\mathrm{DG}}(u - u_{\mathrm{DG}}, v) = 0$ for any $v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$.*

Proof. The proof is similar to the one of [10, Theorem 4.9], and follows from the fact that the solution u satisfies $a_{\mathrm{DG}}(u, v) = \int_\Omega f v \,\mathrm{d}\boldsymbol{x}$, for any $v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$. To prove this identity, we first note that, for any $u \in N^2_{-1-\boldsymbol{\beta}}(\Omega; \mathcal{C}, \mathcal{E}_D)$ and $v \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$, there holds the Green's formula

$$\int_K v \Delta u \;\mathrm{d}\boldsymbol{x} = \int_K \nabla u \cdot \nabla v \;\mathrm{d}\boldsymbol{x} - \int_{\partial K} (\nabla u \cdot \boldsymbol{n}_K) v \;\mathrm{d}s, \qquad \forall\, K \in \mathcal{M}, \tag{4.6}$$

where in the case $\partial K \cap \partial\Omega \neq \emptyset$, the boundary term has to be understood as a pairing in $L^1(\partial K) \times L^\infty(\partial K)$. The formula (4.6) is proved along the lines of [10, Lemma 4.8] with the aid of Lemma 4.1 with $q = 1$. Employing (4.6), the term $\int_\Omega \boldsymbol{A} \nabla u \cdot \nabla_h v \;\mathrm{d}\boldsymbol{x}$ can be integrated by parts on each element, thereby revealing that $-\int_\Omega v \Delta u \;\mathrm{d}\boldsymbol{x} = \int_\Omega f v \,\mathrm{d}\boldsymbol{x}$. Here, the remaining boundary and inter-element flux terms vanish since $[\![u]\!]|_f = 0$ along all $f \in \mathcal{F}_D(\mathcal{M}) \cup \mathcal{F}_I(\mathcal{M})$, and that $[\![\nabla u]\!]|_f = 0$ on all interior faces $f \in \mathcal{F}_I(\mathcal{M})$. The proof of the latter identity is similar to the proof of [10, Lemma 4.7]. $\square$

Finally, the following proposition results from minor modifications of the proofs of the corresponding stability results presented in [10, Theorem 4.4].

Proposition 4.3. *For any σ-geometric mesh $\mathcal{M}$ and degree vector $\boldsymbol{p}(\mathcal{M})$, the bilinear form a_{DG} is continuous and coercive on $V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$: there exist constants $0 < C_2 \leq C_1 < \infty$ independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors such that $|a_{\mathrm{DG}}(v,w)| \leq C_1 \|v\|_{\mathrm{DG}} \|w\|_{\mathrm{DG}}$ for all $v,w \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$, and such that, for $\gamma > 0$ sufficiently large independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors we have $a_{\mathrm{DG}}(v,v) \geq C_2 \|v\|_{\mathrm{DG}}^2$ for all $v \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$. In particular, there exists a unique solution u_{DG} of* (4.2) *(unique up to constants in the pure Neumann case).*

5. Error analysis and exponential convergence

We begin the error analysis by choosing the approximation operators for elements $\mathfrak{O}_\sigma^\ell$ and $\mathfrak{T}_\sigma^\ell$, respectively, and by establishing some of their properties. Then we derive generic errror estimates along the lines of those presented in [11]. Finally, we state our main result: an exponential convergence bound in the dG-norm $\|\cdot\|_{\mathrm{DG}}$ for solutions $u \in B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ as in Proposition 2.3.

5.1. **The elemental approximation operators.** Let u be the solution of (1.1)–(1.3). In this section, we specify a polynomial approximation operator $\Pi u \in V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$. Since functions in $V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})$ are discontinuous, we choose Πu elementwise as $(\Pi u)|_K = \Pi_K u|_K$ for any $K \in \mathcal{M}$.

5.1.1. *L^2-projection in one dimension.* For a generic, bounded interval $I = (a,b)$, we write π_p for the L^2-projection into the space $\mathbb{P}_p(I)$ of degree at most $p \geq 0$ on I (for simplicity we do not explicitly indicate the dependence of π_p on I; this dependence will always be clear from the context). For the purpose of scaling arguments, we further denote by $\widehat{\pi}_p$ the L^2-projection on the reference interval $\widehat{I} = (-1,1)$. The following (p-dependent) stability properties with respect to Sobolev semi-norms will play a crucial role in our analysis.

Lemma 5.1. *Let $I = (a,b)$ be an interval of size $h = b-a$, $p \geq 0$, and $v \in H^j(I)$ for $j \in \mathbb{N}_0$. Then, for every $p \geq j$, there holds the bound*

$$\|(\pi_p u)^{(j)}\|_{L^2(I)} \leq C p^{2j} \|u^{(j)}\|_{L^2(I)} \,, \tag{5.1}$$

where $C > 0$ is a constant depending only on j.

Proof. The L^2-stability of π_p on I, that is the case $j = 0$, is clear and the inequality holds with constant $C = 1$. Next, consider the case $j \geq 1$. Upon scaling it is sufficient to consider the interval $\widehat{I} = (-1,1)$. For $p \geq j$, it holds that $(\widehat{\pi}_p(u))^{(j)} \in \mathbb{P}_{p-j}(\widehat{I})$, and, for the L^2-projections $\widehat{\pi}_{j-1}(u) \in \mathbb{P}_{j-1}(\widehat{I})$, $j = 1,2,\ldots$, we have that

$$\|(\widehat{\pi}_p(u))^{(j)}\|_{L^2(\widehat{I})} = \|(\widehat{\pi}_p(u) - \widehat{\pi}_{j-1}(u))^{(j)}\|_{L^2(\widehat{I})} = \|(\widehat{\pi}_p(u - \widehat{\pi}_{j-1}(u)))^{(j)}\|_{L^2(\widehat{I})} \,.$$

Hence, applying the inverse inequality from [12, Theorem 3.91], yields

$$\|(\widehat{\pi}_p(u))^{(j)}\|_{L^2(\widehat{I})} \leq C_{\mathrm{inv},j} p^{2j} \|u - \widehat{\pi}_{j-1}(u)\|_{L^2(\widehat{I})},$$

and employing a Poincaré-type inequality in $H^j(\widehat{I})/\mathbb{P}_{j-1}(\widehat{I})$, results in

$$\|(\widehat{\pi}_p(u))^{(j)}\|_{L^2(\widehat{I})} \leq C_{inv,j} p^{2j} C_{Poinc,j} \|u^{(j)}\|_{L^2(\widehat{I})} \,.$$

This is the desired estimate. $\square$

5.1.2. *Approximation on $K \in \mathfrak{O}_\sigma^\ell$.* For an interior element $K \in \mathfrak{O}_\sigma^\ell$, we now construct the tensor-product L^2-projection $\Pi_{\boldsymbol{p}_K} u$ as follows. In the setting of (3.2), we write $K = K^\perp \times K^\parallel$, and let $\boldsymbol{p}_K = (p_K^\perp, p_K^\parallel)$. Then we define

$$\begin{aligned} \Pi_{\boldsymbol{p}_K} u|_K &:= \left(\widehat{\Pi}_{\boldsymbol{p}_K}(u \circ \Phi_K)\right) \circ \Phi_K^{-1} \\ &= \left(\widehat{\pi}^{(1)}_{p_K^\perp} \otimes \widehat{\pi}^{(2)}_{p_K^\perp} \otimes \widehat{\pi}^{(3)}_{p_K^\parallel}\right)(u \circ \Phi_K) \circ \Phi_K^{-1} \in \mathbb{Q}_{p_K^\perp}(K^\perp) \otimes \mathbb{Q}_{p_K^\parallel}(K^\parallel), \end{aligned} \tag{5.2}$$

where the one-dimensional L^2-projections act in directions x_1, x_2, and x_3, respectively.

It will be further necessary to distinguish between the perpendicular and parallel projections. To that end, we write

$$\Pi_{\boldsymbol{p}_K} u = (\Pi^{\perp}_{p_K^{\perp}} \otimes \Pi^{\parallel}_{p_K^{\parallel}})u, \qquad K \in \mathfrak{O}^{\ell}_{\sigma}, \tag{5.3}$$

where $(\Pi^{\perp}_{p_K^{\perp}} u)|_K = \left(\widehat{\pi}^{(1)}_{p_K^{\perp}} \otimes \widehat{\pi}^{(2)}_{p_K^{\perp}}\right)(u \circ \Phi_K) \circ \Phi_K^{-1}$, and $\Pi^{\parallel}_{p_K^{\parallel}} u = \widehat{\pi}^{(3)}_{p_K^{\parallel}}(u \circ \Phi_K) \circ \Phi_K^{-1}$.

5.1.3. *A low-order $\mathbb{P}_1$-approximation operator.* We require the following $\mathbb{P}_1$-quasi-interpolation operator considered in [5]. Let $\mathfrak{K} \subset \mathbb{R}^d$ be a bounded, convex polygonal ($d = 2$) or convex polyhedral ($d = 3$) domain which is shape-regular, with diameter $h_{\mathfrak{K}}$, and whose barycenter is

$$\boldsymbol{x}_{\mathfrak{K}} = \frac{1}{|\mathfrak{K}|} \int_{\mathfrak{K}} \boldsymbol{x} \; \mathrm{d}\boldsymbol{x} \in \mathfrak{K}, \tag{5.4}$$

where $|\mathfrak{K}|$ denotes the volume of $\mathfrak{K}$. Then, by definition of $\boldsymbol{x}_{\mathfrak{K}}$,

$$\int_{\mathfrak{K}} (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \; \mathrm{d}\boldsymbol{x} = \boldsymbol{0} \; . \tag{5.5}$$

Define the quasi-interpolation operator $\mathcal{I}_1 : W^{1,1}(\mathfrak{K}) \to \mathbb{P}_1(\mathfrak{K})$ by

$$\mathcal{I}_1 v := \Pi_0 v + (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \cdot \boldsymbol{\Pi}_0(\nabla v), \tag{5.6}$$

where $\mathbb{P}_1(\mathfrak{K})$ denotes the polynomials of total degree at most 1 on $\mathfrak{K}$, and where Π_0 and $\boldsymbol{\Pi}_0$ denote element averages, i.e., the projections onto $\mathbb{P}_0(\mathfrak{K})$ and on $\mathbb{P}_0(\mathfrak{K})^d$, $d = 2, 3$, respectively.

Lemma 5.2. *For the quasi-interpolation operator $\mathcal{I}_1$ defined in* (5.6)*, there holds:*

(1) $\nabla(\mathcal{I}_1 v) \equiv \boldsymbol{\Pi}_0(\nabla v)$ *on $\mathfrak{K}$ for all $v \in W^{1,1}(\mathfrak{K})$.*
(2) $\int_{\mathfrak{K}} (v - \mathcal{I}_1 v) \; \mathrm{d}\boldsymbol{x} = 0$ *and* $\int_{\mathfrak{K}} \nabla(v - \mathcal{I}_1 v) \; \mathrm{d}\boldsymbol{x} = \boldsymbol{0}$ *for all $v \in W^{1,1}(\mathfrak{K})$.*
(3) *For $1 \leq q \leq \infty$, the quasi-interpolant $\mathcal{I}_1$ is $W^{1,q}(\mathfrak{K})$-stable in the following sense:*

$$\forall v \in W^{1,q}(\mathfrak{K}) : \quad \|\nabla(\mathcal{I}_1 v)\|_{L^q(\mathfrak{K})} \leq \|\nabla v\|_{L^q(\mathfrak{K})} \; . \tag{5.7}$$

(4) *For $v \in H^1(\mathfrak{K})$ hold the approximation properties:*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}} \|\nabla v\|_{L^2(\mathfrak{K})}, \qquad \|v - \mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1/2} \|\nabla v\|_{L^2(\mathfrak{K})} \; . \tag{5.8}$$

(5) *If $v \in H^2(\mathfrak{K})$, there holds*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^2 |v|_{H^2(\mathfrak{K})} \; .$$

(6) *Let $d = 2$, and $\boldsymbol{c}$ a corner of $\mathfrak{K}$, and denote $r = r(\boldsymbol{x}) = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{c})$. If $\left\| r^{\beta} \mathsf{D}^{|\boldsymbol{\alpha}|} v \right\|_{L^2(\mathfrak{K})} < \infty$, for any $|\boldsymbol{\alpha}| = 2$ and some $0 < \beta < 1$, then, with an implied constant depending on the shape-regularity of $\mathfrak{K}$, we have*

$$\|v - \mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}} \|\nabla(v - \mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{2-\beta} \sum_{|\boldsymbol{\alpha}|=2} \|r^{\beta} \mathsf{D}^{\boldsymbol{\alpha}} v\|_{L^2(\mathfrak{K})} \; . \tag{5.9}$$

Proof. We prove this lemma item per item.

(1) The first item follows immediately from the definition of $\mathcal{I}_1$ in (5.6).
(2) Moreover, note that

$$v - \mathcal{I}_1 v = (v - \Pi_0 v) - (\boldsymbol{x} - \boldsymbol{x}_{\mathfrak{K}}) \cdot \boldsymbol{\Pi}_0(\nabla v) \; . \tag{5.10}$$

Integrating this identity over $\mathfrak{K}$, the second item follows from property (5.5) and from $\int_{\mathfrak{K}} (v - \Pi_0 v) \; \mathrm{d}\boldsymbol{x} = 0$.

(3) For $1 \leq q < \infty$, the $W^{1,q}(\mathfrak{K})$-stability property results by noticing that $\boldsymbol{\Pi}_0(\nabla v)$ is constant, and from Hölder's inequality:

$$\begin{aligned} \|\nabla(\mathcal{I}_1 v)\|_{L^q(\mathfrak{K})} = \|\boldsymbol{\Pi}_0(\nabla v)\|_{L^q(\mathfrak{K})} &= |\mathfrak{K}|^{1/q} \left| \frac{1}{|\mathfrak{K}|} \int_{\mathfrak{K}} \nabla v \, \mathrm{d}\boldsymbol{x} \right| \\ &\leq |\mathfrak{K}|^{1/q - 1} \|\nabla v\|_{L^q(\mathfrak{K})} \|1\|_{L^{q/(q-1)}(\mathfrak{K})} \leq \|\nabla v\|_{L^q(\mathfrak{K})} \; . \end{aligned}$$

For $q = \infty$ the proof is similar.

(4) To prove the $L^2(\mathfrak{K})$-bound in (5.8), we use (5.6) and (5.7):

$$\begin{aligned}\|v-\mathcal{I}_1 v\|_{L^2(\mathfrak{K})} &\le \|v-\Pi_0 v\|_{L^2(\mathfrak{K})} + \|\Pi_0 v-\mathcal{I}_1 v\|_{L^2(\mathfrak{K})}\\ &= \|v-\Pi_0 v\|_{L^2(\mathfrak{K})} + \|(\boldsymbol{x}-\boldsymbol{x}_{\mathfrak{K}})\cdot\boldsymbol{\Pi}_0(\nabla v)\|_{L^2(\mathfrak{K})}\\ &\lesssim \|v-\Pi_0 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}}\|\nabla v\|_{L^2(\mathfrak{K})}\ .\end{aligned}$$

Furthermore, applying the Poincaré inequality on $H^1(\mathfrak{K})/\mathbb{R}$, there holds $\|v-\Pi_0 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}\|\nabla v\|_{L^2(\mathfrak{K})}$, and thus, the first assertion in (5.8) follows.

In order to prove the second assertion in (5.8), we apply the trace inequality from Lemma 4.1 to the *isotropic* element $\mathfrak{K}$, with $q=2$:

$$\|v-\mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{-1/2}\|v-\mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}}^{1/2}\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})}\ .$$

Taking the gradient of (5.10), we find $\nabla(v-\mathcal{I}_1 v)=\nabla v-\boldsymbol{\Pi}_0(\nabla v)$. We apply the first assertion of (5.8), the triangle inequality, and (5.7) to arrive at $\|v-\mathcal{I}_1 v\|_{L^2(\partial\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1/2}\|\nabla v\|_{L^2(\mathfrak{K})}$.

(5) By item 2 we can employ the Poincaré inequality twice, together with scaling, to obtain

$$\begin{aligned}&\|v-\mathcal{I}_1 v\|_{L^2(\mathfrak{K})} + h_{\mathfrak{K}}\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})}\\ &\qquad\lesssim h_{\mathfrak{K}}\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^2|v-\mathcal{I}_1 v|_{H^2(\mathfrak{K})} = h_{\mathfrak{K}}^2|v|_{H^2(\mathfrak{K})}.\end{aligned}$$

(6) In order to show (5.9), we proceed as in the proof of the previous item and note that

$$\|v-\mathcal{I}_1 v\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})}\ .$$

Thus, it remains to bound $\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})}$. To this end, we apply the first item with the Poincaré inequalities of [9, Proposition 27] or [13, Corollary A.2.11] to find that

$$\|\nabla(v-\mathcal{I}_1 v)\|_{L^2(\mathfrak{K})} = \|\nabla v-\boldsymbol{\Pi}_0(\nabla v)\|_{L^2(\mathfrak{K})} \lesssim h_{\mathfrak{K}}^{1-\beta}\sum_{|\boldsymbol{\alpha}|=2}\left\|r^{\beta}\mathsf{D}^{|\boldsymbol{\alpha}|}v\right\|_{L^2(\mathfrak{K})}\ .$$

This completes the proof. □

5.1.4. *Approximation on* $\mathfrak{T}_\sigma^\ell$. Let $\boldsymbol{e}\in\mathcal{E}$ and consider an element $K=K^\perp\times K^\parallel$ in $\mathfrak{T}_{\boldsymbol{e}}^\ell$ in (3.8). Then we set

$$(\Pi u)|_K = \mathcal{I}_1^\perp\otimes\Pi^\parallel_{p_K^\parallel}u|_K, \qquad K\in\mathfrak{T}_{\boldsymbol{e}}^\ell, \tag{5.11}$$

where $\mathcal{I}_1^\perp$ is the two-dimensional $\mathbb{P}_1$-projector defined in (5.6) and applied in perpendicular direction to $\boldsymbol{e}$ with $\mathfrak{K}=K^\perp$, and $\Pi^\parallel_{p_K^\parallel}$ is the L^2-projection onto polynomials of degree $p_K^\parallel$ in parallel direction to $\boldsymbol{e}$ as in (5.3). Finally, for a corner element $K\in\mathfrak{T}_{\boldsymbol{c}}^\ell$ as in (3.7), we set

$$(\Pi u)|_K := \mathcal{I}_1(u|_K) \tag{5.12}$$

5.1.5. *Tensor-product structure of* Π *on* $\mathfrak{O}_\sigma^\ell$ *and* $\mathfrak{T}_{\mathcal{E}}^\ell$. On elements $K=K^\perp\times K^\parallel$ in $\mathfrak{O}_\sigma^\ell$ and $\mathfrak{T}_{\mathcal{E}}^\ell$, the approximation operator Πu chosen in (5.2), (5.3), and (5.11) has tensor-product structure. In what follows, we shall now simply write

$$(\Pi u)_K = \Pi_K u|_K = \Pi_K^\perp\otimes\Pi_K^\parallel u|_K = (\Pi^\perp\otimes\Pi^\parallel u)|_K, \qquad K\in\mathfrak{O}_\sigma^\ell\cup\mathfrak{T}_{\mathcal{E}}^\ell\ . \tag{5.13}$$

Lemma 5.3. *Let $K\in\mathfrak{O}_\sigma^\ell\cup\mathfrak{T}_{\mathcal{E}}^\ell$. Then Π in (5.13) satisfies:*

(1) *The operator $\Pi_K^\parallel$ is the L^2-projection in edge-parallel direction into polynomials in $\mathbb{P}_{p_K^\parallel}(K^\parallel)$, and $\Pi_K^\perp$ is an approximation operator from $H^1(K^\perp)$ into $\mathbb{Q}_{p_K^\perp}(K^\perp)$ (respectively $\mathbb{P}_{p_K^\perp}(K^\perp)$ in elements K in the terminal layers).*

(2) *The operator $\Pi_K^\perp$ reproduces polynomials in $\mathbb{Q}_{p_K^\perp}(K^\perp)$ (respectively $\mathbb{P}_{p_K^\perp}(K^\perp)$ in elements K in the terminal layers).*

(3) *The operator $\Pi_K^\perp$ satisfies the approximation property:*

$$\|v-\Pi_{K^\perp}^\perp v\|_{L^2(\partial K^\perp)}^2 \lesssim h_{K^\perp}\|\mathsf{D}_\perp v\|_{L^2(K^\perp)}^2, \qquad v\in H^1(K^\perp), \tag{5.14}$$

Proof. The first two properties follow by construction. The trace approximation bound (5.14) is a standard result for the two-dimensional L^2-projection $\Pi_K^\perp=\Pi^\perp_{p_K^\perp}$ in (5.3). For $\Pi_K^\perp=\mathcal{I}_1^\perp$ in (5.11) this follows from (5.8) in Lemma 5.2. □

If now u is the solution of (1.1)–(1.3), and Π the tensor product projection introduced in (5.13), we shall always denote by η the approximation error

$$\eta|_K := u|_K - (\Pi u)|_K, \qquad K \in \mathcal{M}. \tag{5.15}$$

In accordance with (5.13), we also set

$$\eta^\perp|_K := u|_K - (\Pi^\perp u)|_K, \qquad \eta^\|\,|_K := u|_K - (\Pi^\| u)|_K, \qquad K \in \mathfrak{O}^\ell_\sigma \cup \mathfrak{T}^\ell_{\mathcal{E}}\,. \tag{5.16}$$

For $K \in \mathfrak{O}^\ell_\sigma \cup \mathfrak{T}^\ell_{\mathcal{E}}$, we shall further split $\eta|_K$ into

$$\eta|_K = (u|_K - (\Pi^\| u)|_K) + \Pi^\|_K(u|_K - (\Pi^\perp u)|_K) = \eta^\||_K + \Pi^\|_K \eta^\perp|_K. \tag{5.17}$$

The stability of the L^2-projection in (5.1), and the commutativity of the L^2-projectors in perpendicular and parallel direction yields

$$\|\mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\|}_\| \eta\|^2_{L^2(K)} \lesssim (p^\|_K)^{4\alpha^\|} \left(\|\mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\|}_\| \eta^\|\|^2_{L^2(K)} + \|\mathsf{D}^{\boldsymbol{\alpha}^\perp}_\perp \mathsf{D}^{\alpha^\|}_\| \eta^\perp\|^2_{L^2(K)}\right), \tag{5.18}$$

for any $\boldsymbol{\alpha}^\perp \in \mathbb{N}^2_0$, and $0 \le \alpha^\| \le 2$, with "$\lesssim$" uniform in the aspect ratio of K.

5.2. **An anisotropic jump estimate.** The following bound is crucial for controlling the consistency error in anisotropic elements in the terminal layers near Neumann edges.

Proposition 5.4. *Consider an interior face $f = (\partial K_1 \cap \partial K_2)^\circ$ that is parallel to the closest edge $\boldsymbol{e} \in \mathcal{E}$, and which is shared by two axiparallel elements $K_1 = K_1^\perp \times K^\|$ and $K_2 = K_2^\perp \times K^\|$ of possibly high aspect ratios ${}^{(b^\| - a^\|)}/_{h_{K_i^\perp}}$, where $K^\| = (a^\|, b^\|)$, and $h^\|_K = b^\| - a^\|$ denotes the element size in edge-parallel direction, and $K_1^\perp$ and $K_2^\perp$ are two neighboring (but possibly non-matching) rectangles in edge-perpendicular direction such that the bounded variation property* (3.12) *holds. Moreover, for $u \in H^1((\overline{K}_1 \cup \overline{K}_2)^\circ)$, we let $\Pi_{K_i} = \Pi^\perp_{K_i} \otimes \Pi^\|_{K_i}$ be a tensor-product quasi-interpolation operator as in* (5.13) *satisfying properties (1)–(3) in Lemma 5.3 for $i = 1, 2$. Then for η, $\eta^\perp$ and $\eta^\|$ as in* (5.15), (5.16)*, there holds*

$$h_f^{-1}\|[\![\eta]\!]\|^2_{L^2(f)} \lesssim \|\mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_1)} + \|\mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_2)}. \tag{5.19}$$

Proof. Since $\Pi^\perp_{K_i}$ reproduces polynomials in perpendicular direction, we see that

$$\eta^\perp - \Pi^\perp_{K_i}\eta^\perp = (u - \Pi^\perp_{K_i} u) - \Pi^\perp_{K_i}(u - \Pi^\perp_{K_i} u) = u - \Pi^\perp_{K_i} u = \eta^\perp,$$

on K_i, $i = 1, 2$. Since $[\![\eta]\!] = [\![\Pi u]\!]$ and $\Pi^\||_{K_1} u|_{K_1} = \Pi^\|_{K_2} u|_{K_2}$ on f, we obtain

$$\begin{aligned}
\|[\![\eta]\!]\|^2_{L^2(f)} &= \int_f \left(\Pi^\perp_{K_1} \otimes \Pi^\|_{K_1} u|_{K_1} - \Pi^\perp_{K_2} \otimes \Pi^\|_{K_2} u|_{K_2}\right)^2 ds \\
&= \int_f \left((\Pi^\perp_{K_1} \otimes \Pi^\|_{K_1} u|_{K_1} - \Pi^\|_{K_1} u|_{K_1}) - (\Pi^\perp_{K_2} \otimes \Pi^\|_{K_2} u|_{K_2} - \Pi^\|_{K_2} u|_{K_2})\right)^2 ds \\
&\lesssim \int_f \left(\Pi^\|_{K_1} \eta^\perp|_{K_1}\right)^2 ds + \int_f \left(\Pi^\|_{K_2} \eta^\perp|_{K_2}\right)^2 ds.
\end{aligned}$$

Applying (5.14) in perpendicular direction and the L^2-stability of the L^2-projection $\Pi^\|_{K_i}$ yields

$$\begin{aligned}
\|[\![\eta]\!]\|^2_{L^2(f)} &\lesssim h_{K_1^\perp}\|\Pi^\| \mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_1)} + h_{K_2^\perp}\|\Pi^\| \mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_2)} \\
&\lesssim h_{K_1^\perp}\|\mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_1)} + h_{K_2^\perp}\|\mathsf{D}_\perp \eta^\perp\|^2_{L^2(K_2)}\,.
\end{aligned}$$

By the definition of h_f, the bounded variation property (3.12) and the equivalence of $h_{K_1} \simeq h_{K_2}$, we remark that $h_f \simeq h^\perp_{K_1,f} \simeq h^\perp_{K_2,f} \simeq h_{K_1^\perp} \simeq h_{K_2^\perp}$, which implies (5.19). $\square$

5.3. **Error estimates.** To derive error estimates, we proceed in a standard way and split the discretization error $e_{\rm DG} = u - u_{\rm DG}$ into two parts η and ξ, $e_{\rm DG} = \eta + \xi$, with

$$\eta|_K = (u - \Pi u)|_K \qquad \xi|_K = (\Pi u - u_{\rm DG})|_K, \qquad K \in \mathcal{M}_\sigma^{(\ell)}. \tag{5.20}$$

Here, $\Pi u \in V(\mathcal{M}, \boldsymbol{\Phi}, \boldsymbol{p})$ is a polynomial approximation operator as in Section 5.1.

In acccordance with the partition of $\mathcal{M}_\sigma^{(\ell)}$ in (3.3), (3.7), and (3.8), we define the error terms

$$\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta] := \sum_{K\in\mathfrak{O}_\sigma^\ell} T_{\mathfrak{O}}^K[\eta], \qquad \Upsilon_{\mathfrak{T}_{\boldsymbol{e},i}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{e}}^\ell} T_{\boldsymbol{e},i}^K[\eta], \qquad \Upsilon_{\mathfrak{T}_{\boldsymbol{c}}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{c}}^\ell} T_{\boldsymbol{c}}^K[\eta], \tag{5.21}$$

for $i = 1, 2$, where

$$T_{\mathfrak{O}}^K[\eta] := (h_K^{\|})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + (h_K^{\perp})^2\|\mathsf{D}_\perp^2\eta\|_{L^2(K)}^2 + (h_K^{\|})^2\|\mathsf{D}_\|^2\eta\|_{L^2(K)}^2, \tag{5.22}$$

$$T_{\boldsymbol{e},1}^K[\eta] := (h_K^{\|})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + (h_K^{\|})^2\|\mathsf{D}_\|^2\eta\|_{L^2(K)}^2, \tag{5.23}$$

$$T_{\boldsymbol{e},2}^K[\eta] := |K|^{-1}(h_K^{\perp})^2\|\mathsf{D}_\perp^2\eta\|_{L^1(K)}^2, \tag{5.24}$$

$$T_{\boldsymbol{c}}^K[\eta] := h_K^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 + h_K^{-1}|\eta|_{W^{2,1}(K)}^2\ . \tag{5.25}$$

In addition, for a Dirichlet edge $\boldsymbol{e} \in \mathcal{E}_D$, we set

$$\Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta] := \sum_{K\in\mathfrak{T}_{\boldsymbol{e}}^\ell} T_{\boldsymbol{e},D}^K[\eta], \qquad T_{\boldsymbol{e},D}^K[\eta] = (h_K^{\perp})^{-2}\|\eta\|_{L^2(K)}^2. \tag{5.26}$$

By property (5.18) there holds:

$$\begin{aligned}\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta] &\lesssim \boldsymbol{p}_{\max}^8\big(\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\|]\big), \qquad \Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta] \lesssim \boldsymbol{p}_{\max}^8\big(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\|]\big),\\ \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta] &\lesssim \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\|].\end{aligned} \tag{5.27}$$

Theorem 5.5. *Let $u \in N_{-1-\boldsymbol{b}}^2(\Omega, \mathcal{C}, \mathcal{E}_D)$ be the solution of* (1.1)–(1.3), *and let $u_{\rm DG}$ be the DG approximation obtained from* (4.2) *with a sufficiently large penalty parameter $\gamma > 0$ in the dG space V_σ^ℓ in* (3.10), *respectively in $V_{\sigma,\mathfrak{s}}^\ell$ in* (3.11), *for a σ-geometric axiparallel mesh $\mathcal{M}$. Let $\eta = u - \Pi u$ with Π chosen in Section 5.1. Then for the approximation errors in* (5.15), (5.16) *there holds the error bound*

$$\begin{aligned}\|u - u_{\rm DG}\|_{\rm DG}^2 \le C\boldsymbol{p}_{\max}^{12}\Bigg(&\Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{O}_\sigma^\ell}[\eta^\|] + \sum_{\boldsymbol{e}\in\mathcal{E}}\big(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^\ell}[\eta^\|]\big)\\ &+ \sum_{\boldsymbol{e}\in\mathcal{E}}\Upsilon_{\mathfrak{T}_{\boldsymbol{e},2}^\ell}[\eta] + \sum_{\boldsymbol{c}\in\mathcal{C}}\Upsilon_{\mathfrak{T}_{\boldsymbol{c}}^\ell}[\eta] + \sum_{\boldsymbol{e}\in\mathcal{E}_D}\big(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\perp] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^\ell}[\eta^\|]\big)\Bigg).\end{aligned} \tag{5.28}$$

The constant $C > 0$ is independent of the refinement level ℓ, the local mesh sizes and the local polynomial degree vectors.

Proof. Starting from (5.20), the Galerkin orthogonality in Proposition 4.2, implies that $a_{\rm DG}(\xi, \xi) = -a_{\rm DG}(\eta, \xi)$. Hence, by the coercivity of $a_{\rm DG}$ in Proposition 4.3, we arrive at

$$\|\xi\|_{\rm DG}^2 \lesssim -a_{\rm DG}(\eta, \xi) =: T_1 + T_2, \tag{5.29}$$

where

$$T_1 = \sum_{K\in\mathcal{M}}\int_K \nabla_h\eta\cdot\nabla_h\,\xi\,\mathrm{d}\boldsymbol{x} + \theta\int_{\mathcal{F}_I(\mathcal{M})\cup\mathcal{F}_D(\mathcal{M})}\langle\!\langle\nabla_h\xi\rangle\!\rangle\cdot[\![\eta]\!]\ \mathrm{d}s + \gamma\int_{\mathcal{F}_I(\mathcal{M})\cup\mathcal{F}_D(\mathcal{M})}\mathrm{j}\,[\![\eta]\!]\cdot[\![\xi]\!]\ \mathrm{d}s,$$

and

$$T_2 = -\int_{\mathcal{F}_I(\mathcal{M})\cup\mathcal{F}_D(\mathcal{M})}\langle\!\langle\nabla_h\eta\rangle\!\rangle\cdot[\![\xi]\!]\,\mathrm{d}s\ .$$

The first term is bounded using the Cauchy-Schwarz inequality:

$$
\begin{aligned}
|T_1| \lesssim \boldsymbol{p}_{\max} & \left(\|\nabla_h \eta\|^2_{L^2(\Omega)} + \left\| \mathtt{h}^{-1/2} [\![\eta]\!] \right\|^2_{L^2(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))} \right)^{1/2} \\
& \times \left(\|\nabla_h \xi\|^2_{L^2(\Omega)} + \left\| \mathsf{j}^{-1/2} \langle\!\langle \nabla_h \xi \rangle\!\rangle \right\|^2_{L^2(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))} + \left\| \mathsf{j}^{1/2} [\![\xi]\!] \right\|^2_{L^2(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))} \right)^{1/2} .
\end{aligned}
$$

Estimating the term involving $\langle\!\langle \nabla_h \xi \rangle\!\rangle$ as in the proof of [10, Theorem 4.10], with the aid of [10, Lemma 4.3a)], we obtain

$$
|T_1| \lesssim \boldsymbol{p}_{\max} \left(\|\nabla_h \eta\|^2_{L^2(\Omega)} + \left\| \mathtt{h}^{-1/2} [\![\eta]\!] \right\|^2_{L^2(\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M}))} \right)^{1/2} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} . \tag{5.30}
$$

Next, we bound T_2: There holds

$$
\begin{aligned}
|T_2| &= \sum_{f \in \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \int_f |\langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f| |[\![\xi]\!]| \, \mathrm{d}s \\
&\lesssim \sum_{f \in \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} \|\mathsf{j}^{-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f \|_{L^1(f)} \|\mathsf{j}^{1/2} [\![\xi]\!] \|_{L^\infty(f)},
\end{aligned}
$$

where $\boldsymbol{n}_f$ is an orthonormal vector on f pointing in a preset direction. Therefore, using [10, Lemma 4.3b)], it follows that

$$
\begin{aligned}
|T_2| &\lesssim \boldsymbol{p}^2_{\max} \sum_{f \in \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} |f|^{-1/2} \|\mathsf{j}^{-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f \|_{L^1(f)} \|\mathsf{j}^{1/2} [\![\xi]\!] \|_{L^2(f)} \\
&\lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \left(\sum_{f \in \mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})} |f|^{-1} \|\mathsf{j}^{-1/2} \langle\!\langle \nabla_h \eta \rangle\!\rangle \cdot \boldsymbol{n}_f \|^2_{L^1(f)} \right)^{1/2} \\
&\lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \left(\sum_{K \in \mathcal{M}} \sum_{f \in (\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})) \cap \mathcal{F}_K} |f|^{-1} h^{\perp}_{K,f} \|\nabla_h \eta \cdot \boldsymbol{n}_K \|^2_{L^1(f)} \right)^{1/2} .
\end{aligned}
$$

Since $|\nabla \eta \cdot \boldsymbol{n}_K| = |\partial_{K,f,\perp} \eta|$ on $f \in \mathcal{F}_K$, and $|K| \simeq |f| h^{\perp}_{K,f}$, applying the anisotropic trace inequality (4.5) with $q = 1$ yields

$$
|T_2| \lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \left(\sum_{K \in \mathcal{M}} \sum_{f \in (\mathcal{F}_I(\mathcal{M}) \cup \mathcal{F}_D(\mathcal{M})) \cap \mathcal{F}_K} |K|^{-1} \left(\|\nabla \eta\|^2_{L^1(K)} + (h^{\perp}_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} \right) \right)^{1/2} .
$$

Using that $\|\nabla \eta\|_{L^1(K)} \leq |K|^{1/2} \|\nabla \eta\|_{L^2(K)}$ by Hölder's inequality, we conclude that

$$
|K|^{-1} \|\nabla \eta\|^2_{L^1(K)} \leq \|\nabla \eta\|^2_{L^2(K)}.
$$

Since all elements K are axiparallel hexahedra, there are only two cases, $f \parallel \boldsymbol{e}$ and $f \perp \boldsymbol{e}$, where $\boldsymbol{e}$ is the edge nearest to $f \in \mathcal{F}_K$. In the former case, there holds $(h^{\perp}_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} = (h^{\perp}_K)^2 \|\mathsf{D}^2_{\perp} \eta\|^2_{L^1(K)}$, and in the latter $(h^{\perp}_{K,f})^2 \|\partial^2_{K,f,\perp} \eta\|^2_{L^1(K)} = (h^{\parallel}_K)^2 \|\mathsf{D}^2_{\parallel} \eta\|^2_{L^1(K)}$. Therefore,

$$
|T_2| \lesssim \boldsymbol{p}^2_{\max} |\!|\!| \xi |\!|\!|_{\mathrm{DG}} \left(\sum_{K \in \mathcal{M}} \left(\|\nabla \eta\|^2_{L^2(K)} + |K|^{-1} (h^{\perp}_K)^2 \|\mathsf{D}^2_{\perp} \eta\|^2_{L^1(K)} + |K|^{-1} (h^{\parallel}_K)^2 \|\mathsf{D}^2_{\parallel} \eta\|^2_{L^1(K)} \right) \right)^{1/2} .
$$

Combining this estimate with (5.29) and (5.30), dividing the resulting inequality by $\|\xi\|_{\mathrm{DG}}$ and squaring results in

$$\|\xi\|_{\mathrm{DG}}^2 \lesssim \boldsymbol{p}_{\max}^4 \Bigg(\|\nabla_h \eta\|_{L^2(\Omega)}^2 + \sum_{f\in\mathcal{F}_I\cup\mathcal{F}_D} \mathtt{h}_f^{-1} \, \|[\![\eta]\!]\|_{L^2(f)}^2 + \sum_{K\in\mathcal{M}} |K|^{-1} \left((h_K^{\perp})^2 \|\mathsf{D}_{\perp}^2 \eta\|_{L^1(K)}^2 + (h_K^{\|})^2 \|\mathsf{D}_{\|}^2 \eta\|_{L^1(K)}^2 \right) \Bigg) .$$

Noticing that $\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 \leq 2\|\eta\|_{\mathrm{DG}}^2 + 2\|\xi\|_{\mathrm{DG}}^2$, leads to

$$\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 \lesssim \boldsymbol{p}_{\max}^4 \Bigg(\|\nabla_h \eta\|_{L^2(\Omega)}^2 + \sum_{f\in\mathcal{F}_I\cup\mathcal{F}_D} \mathtt{h}_f^{-1} \, \|[\![\eta]\!]\|_{L^2(f)}^2 + \sum_{K\in\mathcal{M}} |K|^{-1} \left((h_K^{\perp})^2 \|\mathsf{D}_{\perp}^2 \eta\|_{L^1(K)}^2 + (h_K^{\|})^2 \|\mathsf{D}_{\|}^2 \eta\|_{L^1(K)}^2 \right) \Bigg) . \tag{5.31}$$

It remains to bound the jump of η. To this end, we distinguish several cases:

- If $f \perp e$ is an *interior face perpendicular to the closest edge* $\boldsymbol{e}\in\mathcal{E}$, shared by two elements K_1 and K_2, with $\mathtt{h}_f \simeq h_{K_1,f}^{\perp} \simeq h_{K_2,f}^{\perp} \simeq h_{K_1}^{\|} \simeq h_{K_2}^{\|}$, we use the trace estimate (4.5) with $q=2$ to obtain
$$\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 \lesssim \sum_{i=1}^{2} \left((h_{K_i}^{\|})^{-2} \|\eta\|_{L^2(K_i)}^2 + \|\nabla\eta\|_{L^2(K_i)}^2 \right) .$$
- For the jumps over *interior anisotropic faces $f\|e$ which are parallel to* $\boldsymbol{e}\in\mathcal{E}$ (and which are shared by two neighboring elements K_1 and K_2), we apply the anisotropic jump estimate in Proposition 5.4, and see that
$$\begin{aligned}\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 &\lesssim \|\mathsf{D}_{\perp}\eta^{\perp}\|_{L^2(K_1)}^2 + \|\mathsf{D}_{\perp}\eta^{\perp}\|_{L^2(K_2)}^2 \\ &\lesssim \|\nabla\eta^{\perp}\|_{L^2(K_1)}^2 + \|\nabla\eta^{\perp}\|_{L^2(K_2)}^2 .\end{aligned}$$
- Finally, for jumps which abut at a *Dirichlet boundary face*, we apply the trace estimate (4.5) with $q=2$ to obtain, for any $f\in\mathcal{F}_D(\mathcal{M})\cap\mathcal{F}_K$,
$$\begin{aligned}\mathtt{h}_f^{-1}\|[\![\eta]\!]\|_{L^2(f)}^2 \simeq (h_K^{\perp})^{-1}\|\eta\|_{L^2(f)}^2 &\lesssim (h_K^{\perp})^{-2}\|\eta\|_{L^2(K)}^2 + \|\mathsf{D}_{\perp}\eta\|_{L^2(K)}^2 \\ &\lesssim (h_K^{\perp})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 .\end{aligned}$$

Inserting these bounds into (5.31) results in

$$\begin{aligned}\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 \lesssim \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}} \Big(&(h_K^{\|})^{-2}\|\eta\|_{L^2(K)}^2 + \|\nabla\eta\|_{L^2(K)}^2 \\ &+|K|^{-1}(h_K^{\perp})^2\|\mathsf{D}_{\perp}^2\eta\|_{L^1(K)}^2 + |K|^{-1}(h_K^{\|})^2\|\mathsf{D}_{\|}^2\eta\|_{L^1(K)}^2\Big) \\ + \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}\setminus\mathfrak{T}_{\mathcal{C}}^{\ell}} &\|\nabla\eta^{\perp}\|_{L^2(K)}^2 + \boldsymbol{p}_{\max}^4 \sum_{\boldsymbol{e}\in\mathcal{E}_D} \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^{\ell}}[\eta] .\end{aligned}$$

For $K\in\mathfrak{O}_{\sigma}^{\ell}$ and for $K\in\mathfrak{T}_{\boldsymbol{e}}^{\ell}$, $\boldsymbol{e}\in\mathcal{E}$, we estimate the $L^1(K)$-norms of $\mathsf{D}_{\perp}^2\eta$ and $\mathsf{D}_{\|}^2\eta$ (the latter only for $K\in\mathfrak{O}_{\sigma}^{\ell}$) by their $L^2(K)$-norms using Hölder's inequality. Moreover, noting that elements in $\mathfrak{T}_{\mathcal{C}}^{\ell}$ are isotropic with $h_K \simeq h_K^{\perp} \simeq h_K^{\|}$ and $|K| \simeq h_K^3$, yields

$$\begin{aligned}\|u-u_{\mathrm{DG}}\|_{\mathrm{DG}}^2 &\lesssim \boldsymbol{p}_{\max}^4 \Upsilon_{\mathfrak{O}_{\sigma}^{\ell}}[\eta] + \boldsymbol{p}_{\max}^4 \sum_{\boldsymbol{e}\in\mathcal{E}} \left(\Upsilon_{\mathfrak{T}_{\boldsymbol{e},1}^{\ell}}[\eta] + \Upsilon_{\mathfrak{T}_{\boldsymbol{e},2}^{\ell}}[\eta] \right) \\ &+ \boldsymbol{p}_{\max}^4 \sum_{c\in\mathcal{C}} \Upsilon_{\mathfrak{T}_{c}^{\ell}}[\eta] + \boldsymbol{p}_{\max}^4 \sum_{K\in\mathcal{M}\setminus\mathfrak{T}_{\mathcal{C}}^{\ell}} \|\nabla\eta^{\perp}\|_{L^2(K)}^2 + \boldsymbol{p}_{\max}^4 \sum_{\boldsymbol{e}\in\mathcal{E}_D} \Upsilon_{\mathfrak{T}_{\boldsymbol{e},D}^{\ell}}[\eta] .\end{aligned}$$

Employing the splittings (5.27) and recalling that $\mathcal{M} = \mathfrak{O}_{\sigma}^{\ell} \cup \mathfrak{T}_{\mathcal{E}}^{\ell} \cup \mathfrak{T}_{\mathcal{C}}^{\ell}$ implies the assertion. $\square$

5.4. **Exponential convergence.** We are now ready to state the main result of this paper.

Theorem 5.6. *Assume that the right-hand side f of the boundary-value problem* (1.1)–(1.3) *in the axiparallel polyhedron $\Omega \subset \mathbb{R}^3$ belongs to the analytic space $B_{1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$, with a weight vector $\boldsymbol{b}$ satisfying* (2.15) *with $0 < b_{\mathcal{C}}, b_{\mathcal{E}} < 1$ as in Remark 2.4. Then the solution u is in $B_{-1-\boldsymbol{b}}(\Omega;\mathcal{C},\mathcal{E}_D)$ according to Proposition 2.3.*

Furthermore, let $\mathfrak{M}_\sigma = \{\mathcal{M}_\sigma^{(\ell)}\}_{\ell\geq 0}$ be a family of axiparallel σ-geometric meshes as introduced in Section 3.1, and consider the hp-dG discretizations in (4.2) *based on the sequences of approximating subspaces V_σ^ℓ and $V_{\sigma,\mathfrak{s}}^\ell$ defined in* (3.10) *respectively* (3.11)*, with the vector $\boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})$ in* (3.10) *of constant, isotropic and uniform polynomial degrees equal to ℓ for the space V_σ^ℓ, respectively the $\mathfrak{s}$-linear, anisotropic degree distribution $\boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})$ for $V_{\sigma,\mathfrak{s}}^\ell$. All polynomial degrees are assumed greater than or equal to* 3 *in elements not abutting at edges $\boldsymbol{e}$ or corners $\boldsymbol{c}$.*

Then for each $\ell \geq 0$, the hp-dG approximation u_{DG} is well-defined, and as $\ell \to \infty$, the approximate solutions u_{DG} satisfy the error estimate

$$\|u - u_{DG}\|_{\mathrm{DG}} \leq C \exp\left(-b\sqrt[5]{N}\right), \tag{5.32}$$

where $N = \dim(V(\mathcal{M}_\sigma^{(\ell)}, \boldsymbol{\Phi}(\mathcal{M}_\sigma^{(\ell)}), \boldsymbol{p}(\mathcal{M}_\sigma^{(\ell)})))$ denotes the number of degrees of freedom of the discretization for any of the two spaces V_σ^ℓ or $V_{\sigma,\mathfrak{s}}^\ell$.

The constants $b > 0$ and $C > 0$ are independent of N, but depend on σ, $\mathcal{M}^0$, θ, γ, $\min \boldsymbol{b} > 0$, and on which of the polynomial degree vectors $\boldsymbol{p}_1(\mathcal{M}_\sigma^{(\ell)})$ or $\boldsymbol{p}_2(\mathcal{M}_\sigma^{(\ell)})$ are used.

Remark 5.7. In particular, the *hp*-dG interpolant constructed to prove Theorem 5.6 yields an exponential approximation bound of the discretization error in the dG norm as in (5.32) for any $u \in B_{-1-\boldsymbol{b}}(\Omega)$.

The remainder of the paper is devoted to the proof of Theorem 5.6. To this end, we will construct appropriate *hp*-interpolants in Section 6 on interior, edge and corner elements. Furthermore, in Section 7 we show that the individual terms on the right-hand side of (5.28) all converge at an exponential rate. Finally, the proof of Theorem 5.6 will be completed in Section 7.6.

6. Approximation properties of L^2-projections

In this section, we establish some approximation results for L^2-projections as in (5.2), (5.3), for elements $K \in \mathfrak{O}_\sigma^\ell$.

6.1. **One-dimensional projectors and *hp*-approximation results.** As in, e.g., [7, 9, 12], the ensuing exponential convergence proofs are based on projectors $\widehat{\pi}_p$ onto polynomials of degree $p \geq 1$, with error bounds which are explicit in the polynomial degree and the regularity order s on $\widehat{I} = (-1, 1)$.

Lemma 6.1. *For any $3 \leq s \leq p$ and $u \in H^{s+1}(\widehat{I})$, we have*

$$\|u - \widehat{\pi}_p u\|_{H^2(\widehat{I})}^2 \lesssim p^8 \Psi_{p-1,s-1} \|u^{(s+1)}\|_{L^2(\widehat{I})}^2 . \tag{6.1}$$

Proof. From [4, Section 8], it follows that for every $p \geq 3$ there exists a projector $\widehat{\pi}_{p,2} : H^2(\widehat{I}) \to \mathbb{P}_p(\widehat{I})$ that satisfies $(\widehat{\pi}_{p,2}u)^{(2)} = \widehat{\pi}_{p-2}u^{(2)}$ and $(\widehat{\pi}_{p,2})^{(j)}u(\pm 1) = u^{(j)}(\pm 1)$ for $j = 0, 1$. The projector $\widehat{\pi}_{p,2}$ is stable in $H^2(\widehat{I})$. Moreover, for any $3 \leq s \leq p$ and $u \in H^{s+1}(\widehat{I})$, there holds the approximation bound

$$\|u - \widehat{\pi}_{p,2} u\|_{H^2(\widehat{I})}^2 \lesssim \Psi_{p-1,s-1} \|u^{(s+1)}\|_{L^2(\widehat{I})}^2 . \tag{6.2}$$

By the triangle inequality, the fact that $\widehat{\pi}_p$ reproduces polynomials, and by the stability estimate (5.1), we see that

$$\|u - \widehat{\pi}_p u\|_{H^2(\widehat{I})} \leq \|u - \widehat{\pi}_{p,2} u\|_{H^2(\widehat{I})} + \|\widehat{\pi}_p(u - \widehat{\pi}_{p,2}u)\|_{H^2(\widehat{I})} \lesssim p^4 \|u - \widehat{\pi}_{p,2} u\|_{H^2(\widehat{I})}, \tag{6.3}$$

Referring to (6.2) yields the assertion for any $u \in H^{s+1}(\widehat{I})$. $\square$

In the remainder of this subsection, we establish exponential convergence results for a broken hp-interpolant on geometric meshes which will be used later, but which are also of independent interest. To that end, on $\omega = (0, 1)$, we consider a sequence $\{\mathcal{T}_\sigma^\ell\}_{\ell=1}^\infty$ of geometric meshes $\mathcal{T}_\sigma^\ell = \{K_j\}_{j=1}^{\ell+1}$ with $\ell+1$ elements which are geometrically graded towards the origin with grading factor $0 < \sigma < 1$. The elements are given by $K_1 = (0, \sigma^\ell)$ and $K_j = (\sigma^{\ell+2-j}, \sigma^{\ell+1-j})$ for $2 \leq j \leq \ell + 1$. The size of element K_j is given by

$$h_{K_j} = \sigma^{\ell+1-j}(1 - \sigma), \qquad 2 \leq j \leq \ell + 1, \tag{6.4}$$

which implies that there is a constant κ solely depending on σ such that

$$\kappa^{-1} h_{K_j} \leq |x| \leq \kappa h_{K_j}, \qquad x \in K_j,\ 2 \leq j \leq \ell + 1\,. \tag{6.5}$$

For a slope parameter $\mathfrak{s} > 0$, we define on $\mathcal{T}_\sigma^\ell$ a $\mathfrak{s}$-linear polynomial degree vector $\boldsymbol{p}$ of length $\ell+1$ given by $\boldsymbol{p} = (p_1, ..., p_{\ell+1})$, with $p_j = \max\{3, \lceil \mathfrak{s} j \rceil\}$, $j = 1, 2, ..., \ell+1$, and set $|\boldsymbol{p}| = \max_{j=1}^{\ell+1} p_j$. We then consider the one-dimensional hp-version discontinuous finite element space

$$S^{\boldsymbol{p},0}(\omega; \mathcal{T}_\sigma^\ell) = \left\{ u \in L^2(\omega) \,:\, u|_{K_j} \in \mathbb{P}^{p_j}(K_j),\ j = 1, 2, ..., \ell + 1 \right\}. \tag{6.6}$$

Then, we denote by $\pi_{\boldsymbol{p}}$ the L^2-projection onto the space $S^{\boldsymbol{p},0}(\omega; \mathcal{T}_\sigma^\ell)$, defined on each element K_j as $(\pi_{\boldsymbol{p},0} u)|_{K_j} = \pi_{p_j,0}(u|_{K_j})$, with the elemental L^2-projection π_{p_j} on K_j as introduced in Section 5.1.1. For a function $u : \omega \to \mathbb{R}$, we define the approximation error by $\eta := u - \pi_{\boldsymbol{p}} u$, and introduce the local error norm:

$$T_j[\eta] := h_{K_j}^{-2} \|\eta\|_{L^2(K_j)}^2 + \|\eta'\|_{L^2(K_j)}^2 + h_{K_j}^2 \|\eta''\|_{L^2(K_j)}^2. \tag{6.7}$$

Proposition 6.2. *For a weight $\beta > 0$, let $u : \omega \to \mathbb{R}$ be such that*

$$\||x|^{-1-\beta+s} u^{(s)}\|_{L^2(\omega)} \leq C_u^{s+1} \Gamma(s + 1), \qquad s \geq 2. \tag{6.8}$$

Then for ℓ sufficiently large, we have $\sum_{j=2}^{\ell+1} T_j[\eta] \leq C \exp(-2b\ell)$, with constants $b, C > 0$ which are independent of ℓ.

Proof. Fix an element $K_j \in \mathcal{T}_\sigma^\ell$ for $2 \leq j \leq \ell + 1$. A straightforward scaling argument yields

$$T_j[\eta] \simeq \left(\frac{h_{K_j}}{2}\right)^{-1} \|\widehat{\eta}\|_{H^2(\widehat{K})}^2,$$

where as usual we denote by $\widehat{\eta}$ the pullback of $\eta|_{K_j}$ to the reference interval $\widehat{I} = (-1, 1)$. Therefore the approximation bound (6.1) implies that

$$T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left(\frac{h_{K_j}}{2}\right)^{-1} \Psi_{p_j-1,s_j-1} \|\widehat{u}^{(s_j+1)}\|_{L^2(\widehat{K})}^2,$$

for any $3 \leq s_j \leq p_j$. Scaling the right-hand side above back to element K_j results in

$$T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left(\frac{h_{K_j}}{2}\right)^{2s_j} \Psi_{p_j-1,s_j-1} \|u^{(s_j+1)}\|_{L^2(K_j)}^2. \tag{6.9}$$

Moreover, by the equivalence (6.5),

$$\|u^{(s_j+1)}\|_{L^2(K_j)}^2 \simeq h_{K_j}^{2+2\beta-2(s_j+1)} \||x|^{-1-\beta+(s_j+1)} u^{(s_j+1)}\|_{L^2(K_j)}^2. \tag{6.10}$$

By combining (6.9), (6.10) with (6.8), we find that

$$\begin{aligned} T_j[\eta] &\lesssim |\boldsymbol{p}|^8 h_{K_j}^{2\beta} 2^{-2s_j} \Psi_{p_j-1,s_j-1} \||x|^{-1-\beta+(s_j+1)} u^{(s_j+1)}\|_{L^2(K_j)}^2 \\ &\lesssim |\boldsymbol{p}|^8 h_{K_j}^{2\beta} \left(\frac{C_u}{2}\right)^{2s_j} \Psi_{p_j-1,s_j-1} \Gamma(s_j + 2)^2\,, \end{aligned} \tag{6.11}$$

for any integer index $3 \leq s_j \leq p_j$. An interpolation argument as in [11, Lemma 5.8] shows that the bound (6.11) holds for any real $s_j \in [3, p_j]$.

Next, we sum the bound (6.11) over all layers $2 \leq j \leq \ell+1$. In view of (6.4), we obtain

$$\sum_{j=2}^{\ell+1} T_j[\eta] \lesssim |\boldsymbol{p}|^8 \left(\sum_{j=2}^{\ell+1} \sigma^{2(\ell+1-j)\beta} \min_{s_j \in [3,p_j]} \left[C^{2s_j} \Psi_{p_j-1,s_j-1} \Gamma(s_j+2)^2 \right] \right).$$

In [11, Lemma 5.12], it has been shown that terms of the form as in the bracket on the right-hand side above can be bounded by $C\exp(-2b(\ell+1))$. By possibly increasing the constant $C > 0$ and by reducing the value of b, the algebraic factor $|\boldsymbol{p}|^8$ can be absorbed into the exponential convergence bound. □

Similarly, we obtain the following result.

Proposition 6.3. *For a weight exponent $\beta > 0$, let $u : \omega \to \mathbb{R}$ be such that there exists a constant $C_u > 0$ with*

$$\||x|^{-\beta+s} u^{(s)}\|_{L^2(\omega)} \leq C_u^{s+2}\Gamma(s+2) \qquad \forall\, s \geq 2\,. \tag{6.12}$$

Then there exist constants $b, C > 0$ such that, for every $\ell \geq 2$, we have $\sum_{j=2}^{\ell+1} \|\eta\|^2_{L^2(K_j)} \leq C\exp(-2b\ell)$, with constants $b, C > 0$ which are independent of ℓ.

Proof. We may assume that ℓ is sufficiently large. Fix an element $K_j \in \mathcal{T}_\sigma^\ell$ for $2 \leq j \leq \ell+1$. Scaling gives $\|\eta\|^2_{L^2(K_j)} = {}^{h_{K_j}}/{}_2 \|\widehat{\eta}\|^2_{L^2(\widehat{K})}$. Then, the approximation bound (6.1), a scaling argument, the equivalence (6.5), and the regularity assumption (6.12) yield, for $3 \leq s_j \leq p_j$,

$$\begin{aligned}
\|\eta\|^2_{L^2(K_j)} &\lesssim |\boldsymbol{p}|^8 \Big(\frac{h_{K_j}}{2}\Big) \Psi_{p_j-1,s_j-1} \|\widehat{u}^{(s_j+1)}\|^2_{L^2(\widehat{K})} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \left(\frac{h_{K_j}}{2}\right)^{2s_j+2} \|u^{(s_j+1)}\|^2_{L^2(K_j)} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \left(\frac{h_{K_j}}{2}\right)^{2s_j+2} h_{K_j}^{2\beta-2s_j-2} \||x|^{-\beta+s_j+1} u^{(s_j+1)}\|^2_{L^2(K_j)} \\
&\lesssim |\boldsymbol{p}|^8 \Psi_{p_j-1,s_j-1} \Big(\frac{C_u}{2}\Big)^{2s_j} h_{K_j}^{2\beta} \Gamma(s_j+3)^2\,.
\end{aligned}$$

From here, the desired estimate follows as in the proof of Proposition 6.2. □

6.2. **Approximation properties of L^2-projection on axiparallel hexahedra.** Now we provide approximation properties of the element-average projection (5.2), (5.3). In the setting (5.15), (5.16), we first show the following estimate for $\eta^{\|}$.

Lemma 6.4. *Let K be an axiparallel hexahedron. For $0 \leq |\boldsymbol{\alpha}^\perp|$, $0 \leq \alpha^{\|} \leq 2$, and $3 \leq s_K^{\|} \leq p_K^{\|}$, there holds*

$$\|\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{\mathsf{D}}_\|^{\alpha^\|} \widehat{\eta}^{\|}\|^2_{L^2(\widehat{K})} \lesssim (p_K^{\|})^8\, \Psi_{p_K^{\|}-1, s_K^{\|}-1} (h_K^\perp)^{2|\boldsymbol{\alpha}^\perp|-2} (h_K^{\|})^{2s_K^{\|}+1} \|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp} \mathsf{D}_\|^{s_K^{\|}+1} u\|^2_{L^2(K)}. \tag{6.13}$$

Proof. Note that $\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{\mathsf{D}}_\|^{\alpha^\|} \widehat{\eta}^{\|} = \widehat{\mathsf{D}}_\|^{\alpha^\|} \left((\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{u}) - \widehat{\Pi}^{\|}_{p_K^{\|}} (\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{u}) \right)$. Applying Lemma 6.1 in edge-parallel direction, we obtain $\|\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{\mathsf{D}}_\|^{\alpha^\|} \widehat{\eta}^{\|}\|^2_{L^2(\widehat{K})} \lesssim (p_K^{\|})^4 \Psi_{p_K^{\|}-1, s_K^{\|}-1} \|\widehat{\mathsf{D}}_\perp^{\boldsymbol{\alpha}^\perp} \widehat{\mathsf{D}}_\|^{s_K^{\|}+1} \widehat{u}\|^2_{L^2(\widehat{K})}$. A scaling argument as in [11, Section 5.1.4] implies the bound (6.13). □

Second, we derive the following bound for $\eta^\perp$. To that end, we introduce the tensor-product space $H^2_{\mathrm{mix}}(\widehat{K}) := H^2(\widehat{I}) \otimes H^2(\widehat{I}) \otimes H^2(\widehat{I})$, and endow it with the standard tensor-product norm.

Lemma 6.5. *For an axiparallel element K and $3 \leq s_K^\perp \leq p_K^\perp$, there holds*

$$\|\widehat{\eta}^\perp\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim (p_K^\perp)^{16}\, E^\perp_{p_K^\perp, s_K^\perp}(K), \tag{6.14}$$

with

$$E^\perp_{p_K^\perp, s_K^\perp}(K) = \Psi_{p_K^\perp-1, s_K^\perp-1} \sum_{\substack{s^\perp+1 \leq |\boldsymbol{\alpha}^\perp| \leq s^\perp+3 \\ 0 \leq \alpha^\| \leq 2}} (h_K^\perp)^{2|\boldsymbol{\alpha}^\perp|-2} (h_K^{\|})^{2\alpha^\|-1} \|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp} \mathsf{D}_\|^{\alpha^\|} u\|^2_{L^2(K)}. \tag{6.15}$$

Proof. In view of (5.2), (5.3), we may write

$$\widehat{\eta}^{\perp} = \widehat{u} - \widehat{\pi}^{(1)}_{p_K^{\perp}} \otimes \widehat{\pi}^{(2)}_{p_K^{\perp}} \widehat{u} = (\widehat{u} - \widehat{\pi}^{(1)}_{p_K^{\perp}} \widehat{u}) + \widehat{\pi}^{(1)}_{p_K^{\perp}} \left(\widehat{u} - \widehat{\pi}^{(2)}_{p_K^{\perp}} \widehat{u} \right) .$$

Hence, by the triangle inequality and the stability properties in (5.1), we readily find that

$$\|\widehat{\eta}^{\perp}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim (p_K^{\perp})^8 \left(\sum_{i=1}^{2} \|\widehat{u} - \widehat{\pi}^{(i)}_{p_K^{\perp}} \widehat{u}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \right) .$$

Lemma 6.1 (used in directions x_1 and x_2) now implies

$$\begin{aligned} &\|\widehat{\eta}^{\perp}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \\ &\quad \lesssim (p_K^{\perp})^{16} \Psi_{p_K^{\perp}-1, s_K^{\perp}-1} \Big(\sum_{0 \le \alpha_2^{\perp}, \alpha^{\parallel} \le 2} \|\widehat{\mathsf{D}}^{(s_K^{\perp}+1, \alpha_2^{\perp}, \alpha^{\parallel})} \widehat{u}\|^2_{L^2(\widehat{K})} + \sum_{0 \le \alpha_1^{\perp}, \alpha^{\parallel} \le 2} \|\widehat{\mathsf{D}}^{(\alpha_1^{\perp}, s_K^{\perp}+1, \alpha^{\parallel})} \widehat{u}\|^2_{L^2(\widehat{K})} \Big) . \end{aligned}$$

This bound and a scaling argument as in [11, Section 5.1.4] yield the desired bound. □

Remark 6.6. It is worth pointing out that a tensor-product argument similar to that in the proof of Lemma 6.5 (see also [11, Section 5.2.1]) applied to the tensor-product projector $\Pi_{\boldsymbol{p}_K}$ in (5.2) implies the following bound: for any axiparallel element K and for $\eta = u - \Pi_{\boldsymbol{p}_K} u$, there holds

$$\|\widehat{\eta}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \lesssim |\boldsymbol{p}_K|^{16} \left(E^{\parallel}_{p_K^{\parallel}, s_K^{\parallel}}(K) + E^{\perp}_{p_K^{\perp}, s_K^{\perp}}(K) \right), \tag{6.16}$$

for any $3 \le s_K^{\perp} \le p_K^{\perp}$ and $3 \le s_K^{\parallel} \le p_K^{\parallel}$, with

$$E^{\parallel}_{p_K^{\parallel}, s_K^{\parallel}}(K) = \Psi_{p_K^{\parallel}-1, s_K^{\parallel}-1} \sum_{0 \le \alpha_1^{\perp}, \alpha_2^{\perp} \le 2} (h_K^{\perp})^{2|\boldsymbol{\alpha}^{\perp}|-2} (h_K^{\parallel})^{2 s_K^{\parallel}+1} \|\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}} \mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)}, \tag{6.17}$$

and $E^{\perp}_{p_K^{\perp}, s_K^{\perp}}(K)$ defined in (6.15). Up to the algebraic loss in $|\boldsymbol{p}_K|$, the estimate (6.16) is the same as that in [11, Lemma 5.6] used in the analysis of the pure Dirichlet case. However, in the case of a corner-edge patch involving a Neumann edge, we shall invoke the finer bound in Lemma 6.4.

7. Reference corner-edge patch

According to the construction of the hp-dG spaces provided in Section 3, the geometric edge mesh $\mathcal{M}^{\ell}$ consists of a finite number of physical patches $\{\mathcal{M}_j^{\ell}\}_{j=1}^{J^{\ell}}$. This makes it possible to bound the right-hand side of (5.28) separately on each $\mathcal{M}_j^{\ell}$ by means of a suitable hp-approximation analysis. In addition, noting that each patch $\mathcal{M}_j^{\ell}$ is equivalent (up to isotropic dilation, translation and/or rotation) to one of the reference patches displayed in Figure 1, it is sufficient to limit the proof of the exponential convergence bounds to the reference situations from Figure 1. Indeed, due to the simple structure of the patch mappings, the weighted Sobolev space $N^k_{\boldsymbol{\beta}}(\mathcal{M}_j^{\ell}; \mathcal{C}, \mathcal{E}_D)$, as restricted to a physical patch $\mathcal{M}_j^{\ell}$, can be identified with an equivalent space, which features the same regularity and is equipped with equivalent norms, on one of the reference patches.

7.1. **The setting.** We consider a reference corner-edge patch in $(0,1)^3$ consisting of a single corner $\boldsymbol{c} \in \mathcal{C}$ and a single edge $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ originating from it; see Figure 1 (right) for an illustration. We may assume that $\boldsymbol{c} = (\boldsymbol{0}, 0)$, and $\boldsymbol{e} = \{\boldsymbol{0}\} \times \omega_{\boldsymbol{c}}^{\parallel}$ with $\omega_{\boldsymbol{c}}^{\parallel} = (0,1)$.

Similarly to [11], we now introduce a reference geometric corner-edge mesh $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$. As in [11], $\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$ is built from mesh layers via

$$\widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}} = \bigcup_{j=1}^{\ell+1} \bigcup_{i=1}^{j} \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}, \tag{7.1}$$

where the sets $\widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ stand for layers of elements with identical scaling properties. The decomposition in (7.1) is not a partition, in general: elements may be contained in several layers whose number, however, is uniformly bounded with respect to ℓ.

In (7.1), the index j indicates the number of the geometric mesh layers in edge-parallel direction along the edge $\omega_{\boldsymbol{c}}^{\|}$, whereas the index i indicates the number of mesh layers in direction perpendicular to $\omega_{\boldsymbol{c}}^{\|}$. In agreement with (3.3), (3.7), (3.8), we split $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell}$ into interior elements away from $\boldsymbol{c}$ and $\boldsymbol{e}$, boundary layer elements along $\boldsymbol{e}$ (but away from $\boldsymbol{c}$), and the corner element by setting

$$\widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell} = \widehat{\mathfrak{O}}_{\boldsymbol{ce}}^{\ell} \,\dot{\cup}\, \widehat{\mathfrak{T}}_{\boldsymbol{e}}^{\ell} \,\dot{\cup}\, \widehat{\mathfrak{T}}_{\boldsymbol{c}}^{\ell}, \tag{7.2}$$

where

$$\widehat{\mathfrak{O}}_{\boldsymbol{ce}}^{\ell} := \bigcup_{j=2}^{\ell+1} \bigcup_{i=2}^{j} \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{e}}^{\ell} := \bigcup_{j=2}^{\ell+1} \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{1j}, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{c}}^{\ell} := \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{11}. \tag{7.3}$$

In particular, an interior element $K \in \widehat{\mathfrak{O}}_{\boldsymbol{ce}}^{\ell}$ belongs to $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}$ if it satisfies

$$r_{\boldsymbol{e}}|_K \simeq d_K^{\boldsymbol{e}} \simeq h_K^{\perp} \simeq \sigma^{\ell+1-i}, \qquad r_{\boldsymbol{c}}|_K \simeq d_K^{\boldsymbol{c}} \simeq h_K^{\|} \simeq \sigma^{\ell+1-j}, \qquad 2 \leq i \leq j \leq \ell+1\,. \tag{7.4}$$

Moreover, the terminal layers $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{1j}$ at $\boldsymbol{e} \in \mathcal{E}$ consist of elements $K \in \widehat{\mathfrak{T}}_{\boldsymbol{e}}^{\ell}$ with

$$r_{\boldsymbol{e}}|_K \simeq d_K^{\boldsymbol{e}} \lesssim h_K^{\perp} \simeq \sigma^{\ell}, \qquad r_{\boldsymbol{c}}|_K \simeq d_K^{\boldsymbol{c}} \simeq h_K^{\|} \simeq \sigma^{\ell+1-j}, \qquad 2 \leq j \leq \ell+1, \tag{7.5}$$

Finally, any element in the layer $\widehat{\mathfrak{T}}_{\boldsymbol{c}}^{\ell} = \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{11}$ is isotropic with

$$r_{\boldsymbol{e}}|_K \simeq d_K^{\boldsymbol{e}} \lesssim h_K \simeq \sigma^{\ell}, \qquad r_{\boldsymbol{c}}|_K \simeq d_K^{\boldsymbol{c}} \lesssim h_K \simeq \sigma^{\ell}\,. \tag{7.6}$$

The cardinality of the layers $\mathfrak{L}_{\boldsymbol{ce}}^{ij}$ depends on the implied equivalence constants in (7.4)–(7.6). We emphasize that the ensuing analysis is valid for any choice of these constants (independent of i, j, ℓ). For the reference patch as shown in Figure 1 (right) the sets $\mathfrak{L}_{\boldsymbol{ce}}^{ij}$ are in fact singletons, and any $K \in \mathfrak{L}_{\boldsymbol{ce}}^{1j}$ can be written in the form

$$K_j = K^{\perp} \times K_j^{\|}, \qquad 2 \leq j \leq \ell+1, \tag{7.7}$$

where $K^{\perp} = (0, \sigma^{\ell})^2$, and the sequence $\{K_j^{\|}\}_{j=2}^{\ell+1}$ forms a one-dimensional geometric mesh $\mathcal{T}_{\sigma}^{\ell}$ along the edge $\omega_{\boldsymbol{c}}^{\|} = (0,1)$ as in Section 6.1; moreover, there is a single corner element $K \in \mathfrak{T}_{\boldsymbol{c}}^{\ell}$ that is given by $K = (0, \sigma^{\ell})^3$.

In agreement with the hp-extensions (Ex1)–(Ex4) in [10], we consider $\mathfrak{s}$-linear polynomial degree distributions on $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell}$ that satisfy

$$\forall\, K \in \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}: \qquad \boldsymbol{p}_K = (p_i^{\perp}, p_j^{\|}) \simeq (\max\{\lceil \mathfrak{s} i \rceil, 3\}, \max\{\lceil \mathfrak{s} j \rceil, 3\}), \qquad 1 \leq i \leq j \leq \ell+1\,. \tag{7.8}$$

We note that our hp-approximation analysis below allows for $\max \boldsymbol{p}_K < 3$ in corner elements $K \in \mathfrak{T}_{\mathcal{C}}^{\ell}$.

Let now $\widehat{\Omega}_{\boldsymbol{ce}}^{\ell}$ denote the domain formed by all elements in $\widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell}$:

$$\widehat{\Omega}_{\boldsymbol{ce}}^{\ell} = \Big(\bigcup_{K \in \widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell}} \overline{K} \Big)^{\circ}. \tag{7.9}$$

Analogous to the reference corner-edge patch $\widehat{\Omega}_{\boldsymbol{ce}}^{\ell}$, corresponding to the rightmost display in Fig. 1, we introduce the *reference corner patch* $\widehat{\Omega}_{\boldsymbol{c}}^{\ell}$ and the *reference edge patch* $\widehat{\Omega}_{\boldsymbol{e}}^{\ell}$, which correspond to the leftmost and the middle panel, respectively, in Fig. 1. For the purpose of deriving the ensuing exponential convergence estimates it is important that *the corresponding geometric mesh patches can be characterized as collections of certain elements* $K \in \widehat{\mathcal{M}}_{\boldsymbol{ce}}^{\ell}$: for $\ell \geq 2$, we define with $\widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{ij}$ as in (7.1)

$$\forall \boldsymbol{c} \in \mathcal{C}: \quad \widehat{\mathcal{M}}_{\boldsymbol{c}}^{\ell} = \widehat{\mathfrak{O}}_{\boldsymbol{c}}^{\ell} \cup \widehat{\mathfrak{T}}_{c}^{\ell} \qquad \widehat{\mathfrak{O}}_{\boldsymbol{c}}^{\ell} := \bigcup_{j=2}^{\ell+1} \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{jj}\,, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{c}}^{\ell} := \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{11}\,, \tag{7.10}$$

$$\forall \boldsymbol{e} \in \mathcal{E}: \quad \widehat{\mathcal{M}}_{\boldsymbol{e}}^{\ell} = \widehat{\mathfrak{O}}_{\boldsymbol{e}}^{\ell} \cup \widehat{\mathfrak{T}}_{e}^{\ell} \qquad \widehat{\mathfrak{O}}_{\boldsymbol{e}}^{\ell} := \bigcup_{i=2}^{\ell+1} \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{i,\ell+1}\,, \qquad \widehat{\mathfrak{T}}_{\boldsymbol{e}}^{\ell} := \widehat{\mathfrak{L}}_{\boldsymbol{ce}}^{1,\ell+1}\,. \tag{7.11}$$

We establish exponential convergence of the hp-dGFEM by proving exponential convergence estimates for the consistency bound (5.28) for each of the three canonical geometric mesh patches shown in Fig. 1. We remark that we abuse notation slightly in that the definition of $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{c}}$ and $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$ in (7.10) and (7.11) differs from (7.3); it will be clear from the case discussed which definition is applicable. Due to (7.10) and (7.11), *the required exponential convergence bounds for each of the three basic geometric mesh patches depicted in Fig. 1 will follow from consistency error estimates in patch* $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$ which we therefore now consider next.

For a function $u : \widehat{\Omega}^{\ell}_{\boldsymbol{ce}} \to \mathbb{R}$ (whose regularity will be specified below) and for $K \in \widehat{\mathcal{M}}^{\ell}_{\boldsymbol{ce}}$, we define the elemental approximation operators $(\Pi u)_K = \Pi_K u|_K$ in accordance with the choices in Section 5.1. That is, for interior elements $K \in \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ we select Π_K to be the L^2-projection as in (5.2), (5.3), with the elemental polynomial degrees taken as $p^{\perp}_K = p^{\perp}_i$, $p^{\perp}_K = p^{\parallel}_j$; cp. (7.8). For $K_j \in \mathfrak{T}^{\ell}_{\boldsymbol{e}}$ of the form (7.7), we select Π_{K_j} as in (5.11) with $p^{\parallel}_{K_j} = p^{\parallel}_j$. Finally, for the corner element $K \in \mathfrak{T}^{\ell}_{\boldsymbol{c}}$, we select Π_K in agreement with (5.12). For functions $u : \widehat{\Omega}^{\ell}_{\boldsymbol{c}} \to \mathbb{R}$ and $u : \widehat{\Omega}^{\ell}_{\boldsymbol{e}} \to \mathbb{R}$, we will obtain exponential convergence estimates as direct consequences from the elementwise bounds established in the analysis on $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}} \to \mathbb{R}$.

For the dG approximation errors η, $\eta^{\perp}$, $\eta^{\parallel}$ as in (5.15), (5.16), and in view of the error estimates in Theorem 5.5, we will now bound the contributions $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}$, $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},1}}$, $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},2}}$, and $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{c}}}$, where these terms are defined exactly as in (5.21)–(5.25). If $\boldsymbol{e}$ is a Dirichlet edge, we shall also estimate $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},D}}$ given as in (5.26).

7.2. **Exponential Convergence at Neumann edges.** We shall first consider the case where $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ is a Neumann edge, i.e., $\boldsymbol{e} \in \mathcal{E}_N$. By the regularity property (2.16), the definition of the weighted seminorm (2.11) in the neighbourhood of Neumann edges, and for exponents $b_{\boldsymbol{c}}, b_{\boldsymbol{e}} \in (0,1)$ as in (2.15) and Remark 2.4, the solution u localized in $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$ has finite *corner-edge seminorm* (obtained by localization of (2.11) to $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$)

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\max\{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})}, \qquad k > k_{\boldsymbol{\beta}}, \tag{7.12}$$

with $k_{\boldsymbol{\beta}}$ in (2.12). Under the assumptions on the weights $b_{\boldsymbol{c}}, b_{\boldsymbol{e}}$ in Remark 2.4, we note that, for $\alpha^{\parallel} \geq 0$, the seminorms on the right-hand side of (7.12) take the following forms:

$$\begin{cases} \|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+\alpha^{\parallel}} \mathsf{D}_{\parallel}^{\alpha^{\parallel}} u\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} & |\boldsymbol{\alpha}^{\perp}| = 0, \\ \|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+\alpha^{\parallel}} \mathsf{D}_{\perp} \mathsf{D}_{\parallel}^{\alpha^{\parallel}} u\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} & |\boldsymbol{\alpha}^{\perp}| = 1, \\ \sum_{|\boldsymbol{\alpha}^{\perp}|=k} \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+\alpha^{\parallel}} r_{\boldsymbol{e}}^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|} \mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}} \mathsf{D}_{\parallel}^{\alpha^{\parallel}} u\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} & k = |\boldsymbol{\alpha}^{\perp}| \geq 2\,. \end{cases} \tag{7.13}$$

The corresponding norms $\| \circ \|_{\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})}$ and the weighted spaces $\widehat{N}^m_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ are then defined as in Section 2.2, for $m > k_{\boldsymbol{\beta}}$. For elements $K \in \widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$ we denote by $|\cdot|_{\widehat{N}^k_{-1-\boldsymbol{b}}(K)}$ the restriction of the norm in (7.12) to K, and similarly for the full norm. We say a function $u \in H^1(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ belongs to $B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ if $u \in \widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ for $k > k_{\boldsymbol{\beta}}$ and there is a constant $d_u > 0$ such that

$$\|u\|_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} \leq d_u^{k+1} k!\,, \qquad k > k_{\boldsymbol{\beta}}\,. \tag{7.14}$$

In the corner patch $\widehat{\Omega}^{\ell}_{\boldsymbol{c}}$ and the reference edge patch $\widehat{\Omega}^{\ell}_{\boldsymbol{e}}$ defined in (7.10) and (7.11), respectively, expressions analogous (but simpler) to (7.12) for the respective seminorms result: since $\rho_{\boldsymbol{ce}}|_{\widehat{\Omega}^{\ell}_{\boldsymbol{c}}} = \mathcal{O}(1)$,

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{c}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{c}})} \tag{7.15}$$

(note that in $\widehat{\Omega}^{\ell}_{\boldsymbol{c}}$ the weights are homogeneous as in the Dirichlet case considered in [11]), and in the reference Neumann edge-patch $\widehat{\Omega}^{\ell}_{\boldsymbol{e}}$, we have

$$|u|^2_{\widehat{N}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{e}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{e}}^{\max\{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|,0\}} \mathsf{D}^{\boldsymbol{\alpha}} u \right\|^2_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{e}})}, \tag{7.16}$$

for $k > k_{\boldsymbol{\beta}}$.

7.3. **Exponential Convergence in $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$, $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{e}}$ and $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$.** For Neumann edges $\boldsymbol{e} \in \mathcal{E}_N$ we obtain exponential convergence of all contributions from $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ in the dGFEM consistency error bound (5.28) by an analysis of the corresponding terms in *one reference corner-edge patch* $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$. The general result will then follow upon noting that $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ is obtained by a finite superposition of (scaled and translated versions of) this reference corner-edge patch.

Theorem 7.1. *Let $\boldsymbol{e} \in \mathcal{E}_N$ be a Neumann edge. Let $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as in* (7.12), (7.14), *and let Π_K denote the elemental approximation operators chosen as in Section 5.1. Then for η, $\eta^{\perp}$, $\eta^{|}$ as in* (5.15), (5.16), *there exist constants $b, C > 0$ such that, for ℓ sufficiently large, there holds the exponential convergence estimate*

$$\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] + \Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},1}}[\eta^{\perp}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},1}}[\eta^{\|}] + \Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},2}}[\eta] \leq C \exp(-2b\ell) \,. \tag{7.17}$$

Analogous exponential convergence bounds hold for the consistency terms from $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{c}}$, $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{e}}$.

The proof of the exponential convergence bound (7.17) in Theorem 7.1 will be presented in several steps. The proofs for the bounds on the terms $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{c}}$, $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{e}}$ are analogous (by choosing $h^{\|}_K = \mathcal{O}(1)$ in the proofs which follow) and will not be detailed.

7.3.1. *Exponential convergence of $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}$.* For $\boldsymbol{e} \in \mathcal{E}_N$ and for $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$, we prove exponential convergence of $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}]$ and $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}]$ in (7.17). We begin by recording scalings to the reference cube $\widehat{K} = (-1,1)^3$ of the terms contained in $T^K_{\mathfrak{O}}[v]$ in (5.22).

Lemma 7.2. *For $K \in \widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ and for $v \in H^2(K)$, there holds:*

$$(h^{\|}_K)^{-2}\|v\|^2_{L^2(K)} + \|\mathsf{D}_{\|}v\|^2_{L^2(K)} + (h^{\|}_K)^2\|\mathsf{D}^2_{\|}v\|^2_{L^2(K)} \lesssim (h^{\perp}_K)^2(h^{\|}_K)^{-1}\Big(\sum_{0\leq\alpha^{\|}\leq 2}\|\widehat{\mathsf{D}}^{\alpha^{\|}}_{\|}\widehat{v}\|^2_{L^2(\widehat{K})}\Big), \tag{7.18}$$

as well as

$$(h^{\perp}_K)^{2(|\boldsymbol{\alpha}^{\perp}|-1)}\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}v\|^2_{L^2(K)} \lesssim h^{\|}_K\|\widehat{\mathsf{D}}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\widehat{v}\|^2_{L^2(\widehat{K})}, \qquad |\boldsymbol{\alpha}^{\perp}| = 1, 2\,. \tag{7.19}$$

Proof. These inequalities are an immediate consequence of the scalings in [11, Section 5.1.4]. □

Next, we bound the error term $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}]$ in direction perpendicular to edge $\boldsymbol{e}$.

Proposition 7.3. *Let $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as in* (7.12), (7.14). *Then there exist constants $b, C > 0$ such that for $\ell \geq 2$ holds $\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] \leq C\exp(-2b\ell)$.*

Proof. According to (7.3), we consider $K \in \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ with $2 \leq j \leq \ell+1$ and $2 \leq i \leq j$. The scalings in (7.18), (7.19) and the fact that in $\widehat{\Omega}^{\ell}_{\boldsymbol{ce}}$ holds $h^{\perp}_K \lesssim h^{\|}_K$ allow us to conclude that

$$\begin{aligned} T^K_{\mathfrak{O}}[\eta^{\perp}] &\lesssim \Big((h^{\perp}_K)^2(h^{\|}_K)^{-1} + h^{\|}_K\Big)\|\widehat{\eta}^{\perp}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} = h^{\|}_K\Big(1 + (h^{\perp}_K)^2(h^{\|}_K)^{-2}\Big)\|\widehat{\eta}^{\perp}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})} \\ &\lesssim h^{\|}_K\|\widehat{\eta}^{\perp}\|^2_{H^2_{\mathrm{mix}}(\widehat{K})}\,. \end{aligned}$$

With Lemma 6.5 and (7.8), we obtain

$$T^K_{\mathfrak{O}}[\eta^{\perp}] \lesssim |\boldsymbol{p}_K|^{16}\, h^{\|}_K \Psi_{p^{\perp}_i-1, s^{\perp}_i-1} \sum_{\substack{s^{\perp}+1\leq|\boldsymbol{\alpha}^{\perp}|\leq s^{\perp}+3 \\ 0\leq\alpha^{\|}\leq 2}} (h^{\perp}_K)^{2|\boldsymbol{\alpha}^{\perp}|-2}(h^{\|}_K)^{2\alpha^{\|}-1}\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\mathsf{D}^{\alpha^{\|}}_{\|}u\|^2_{L^2(K)}.$$

Since $K \in \widehat{\mathfrak{L}}^{ij}_{\boldsymbol{ce}}$ with $2 \leq j \leq \ell+1$ and $2 \leq i \leq j$, there hold the equivalences (7.5) on K, and we may insert the appropriate weights according to (7.13) to obtain

$$\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\mathsf{D}^{\alpha^{\|}}_{\|}u\|^2_{L^2(K)} \simeq (d^{\boldsymbol{c}}_K)^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}-2\alpha^{\|}}(d^{\boldsymbol{e}}_K)^{2+2b_{\boldsymbol{e}}-2|\boldsymbol{\alpha}^{\perp}|}\|r^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+\alpha^{\|}}_{\boldsymbol{c}} r^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^{\perp}|}_{\boldsymbol{e}}\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\mathsf{D}^{\alpha^{\|}}_{\|}u\|^2_{L^2(K)}\,.$$

Then we invoke this equivalence and the analytic regularity (7.14) to obtain that there exists a constant $C > 0$ such that for all p_K, $p^{\perp}_i$ and $s^{\perp}_i$ holds

$$T^K_{\mathfrak{O}}[\eta^{\perp}] \lesssim |\boldsymbol{p}_K|^{16}\,\Psi_{p^{\perp}_i-1, s^{\perp}_i-1}(d^{\boldsymbol{c}}_K)^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}(d^{\boldsymbol{e}}_K)^{2b_{\boldsymbol{e}}}C^{2s^{\perp}_i}\Gamma(s^{\perp}_i+6)^2\,. \tag{7.20}$$

Summing (7.20) over all layers in $\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$ in (7.3) with the use of (7.5) results in

$$\Upsilon_{\widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\perp}] \lesssim \boldsymbol{p}_{\max}^{16} \sum_{j=2}^{\ell+1} \sigma^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})(\ell+1-j)} \sum_{i=2}^{j} \sigma^{2b_{e}(\ell+1-i)} \Psi_{p_i^{\perp}-1,s_i^{\perp}-1} C^{2s_i^{\perp}} \Gamma(s_i^{\perp}+6)^2 .$$

By interpolating to real parameters $s_i^{\perp} \in [3, p_i^{\perp}]$ as in [11, Lemma 5.8], this sum is of exactly the same form as $S^{\perp}$ in the proof of [11, Proposition 5.17], and the assertion now follows from the arguments there and after possibly adjusting the constants to absorb the algebraic loss in $\boldsymbol{p}_{\max}$. □

To establish the analog of Proposition 7.3 in edge-parallel direction, we make use of the following estimates.

Lemma 7.4. *Let $K \in \widehat{\mathfrak{O}}^{\ell}_{\boldsymbol{ce}}$, and $3 \leq s_K^{\parallel} \leq p_K^{\parallel}$. Then there holds*

$$(h_K^{\parallel})^{-2} \|\eta^{\parallel}\|^2_{L^2(K)} + \|\mathsf{D}_{\parallel}\eta^{\parallel}\|^2_{L^2(K)} + (h_K^{\parallel})^2 \|\mathsf{D}^2_{\parallel}\eta^{\parallel}\|^2_{L^2(K)} \lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}} |u|^2_{N^{s_K^{\parallel}+1}_{-1-\boldsymbol{b}}(K)}, \tag{7.21}$$

$$\|\mathsf{D}_{\perp}\eta^{\parallel}\|^2_{L^2(K)} \lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}} |u|^2_{N^{s_K^{\parallel}+2}_{-1-\boldsymbol{b}}(K)}, \tag{7.22}$$

as well as

$$(h_K^{\perp})^2 \|\mathsf{D}^2_{\perp}\eta^{\parallel}\|^2_{L^2(K)} \lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{e}})^{2b_{\boldsymbol{e}}} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}} |u|^2_{N^{s_K^{\parallel}+3}_{-1-\boldsymbol{b}}(K)} . \tag{7.23}$$

Proof. We prove (7.21) by bounding the right-hand side in (7.18) with the aid of the approximation property (6.13) (with $|\boldsymbol{\alpha}^{\perp}| = 0$):

$$\begin{aligned}
&(h_K^{\parallel})^{-2} \|\eta^{\parallel}\|^2_{L^2(K)} + \|\mathsf{D}_{\parallel}\eta^{\parallel}\|^2_{L^2(K)} + (h_K^{\parallel})^2 \|\mathsf{D}^2_{\parallel}\eta^{\parallel}\|^2_{L^2(K)} \\
&\quad \lesssim (h_K^{\perp})^2 (h_K^{\parallel})^{-1} \Big(\sum_{0 \leq \alpha^{\parallel} \leq 2} \|\widehat{\mathsf{D}}_{\parallel}^{\alpha^{\parallel}} \widehat{\eta}\|^2_{L^2(\widehat{K})} \Big) \lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (h_K^{\parallel})^{2s_K^{\parallel}} \|\mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} .
\end{aligned}$$

Then, we insert the weight $r_{\boldsymbol{c}}$ by the use of (7.13), (7.4). We find that

$$\|\mathsf{D}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} \simeq (d_K^{\boldsymbol{c}})^{2+2b_{\boldsymbol{c}}-2s_K^{\parallel}-2} \|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+s_K^{\parallel}+1} \mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} \lesssim (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2s_K^{\parallel}} |u|^2_{N^{s_K^{\parallel}+1}_{-1-\boldsymbol{b}}(K)} .$$

Combining the two estimates above shows (7.21).

To establish (7.22), we start from the the right-hand side of (7.19), apply (6.13) (with $|\boldsymbol{\alpha}^{\perp}| = 1$ and $\alpha^{\parallel} = 0$), and insert the appropriate weights employing (7.4). This results in

$$\begin{aligned}
\|\mathsf{D}_{\perp}\eta^{\parallel}\|^2_{L^2(K)} &\lesssim h_K^{\parallel} \|\widehat{\mathsf{D}}_{\perp}\widehat{\eta}^{\parallel}\|^2_{L^2(\widehat{K})} \\
&\lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (h_K^{\parallel})^{2s_K^{\parallel}+2} \|\mathsf{D}_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} \\
&\lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{c}})^{2s_K^{\parallel}+2} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2s_K^{\parallel}-2} \|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+s_K^{\parallel}+1} \mathsf{D}_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} \\
&\lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}} |u|^2_{N^{s_K^{\parallel}+2}_{-1-\boldsymbol{b}}(K)},
\end{aligned}$$

which yields (7.22).

For (7.23), we proceed along the same lines and apply (7.19), (6.13) (with $|\boldsymbol{\alpha}^{\perp}| = 2$ and $\alpha^{\parallel} = 0$), and (7.4). We find that

$$\begin{aligned}
(h_K^{\perp})^2 \|\mathsf{D}^2_{\perp}\eta^{\parallel}\|^2_{L^2(K)} &\lesssim h_K^{\parallel} \|\widehat{\mathsf{D}}^2_{\perp}\widehat{\eta}^{\parallel}\|^2_{L^2(\widehat{K})} \\
&\lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (h_K^{\perp})^2 (h_K^{\parallel})^{2s_K^{\parallel}+2} \|\mathsf{D}^2_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)} \\
&\lesssim (p_K^{\parallel})^8 \Psi_{p_K^{\parallel}-1,s_K^{\parallel}-1} (d_K^{\boldsymbol{e}})^{2b_{\boldsymbol{e}}} (d_K^{\boldsymbol{c}})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}} \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+s_K^{\parallel}+1} r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}} \mathsf{D}^2_{\perp}\mathsf{D}_{\parallel}^{s_K^{\parallel}+1} u\|^2_{L^2(K)},
\end{aligned}$$

which finishes the proof. ☐

We are now ready to bound $\Upsilon_{\widehat{\mathfrak{D}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}]$.

Proposition 7.5. *Let $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as in (7.12), (7.14). Then, there exist $b, C > 0$ such that, for ℓ sufficiently large, there holds $\Upsilon_{\widehat{\mathfrak{D}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}] \leq C\exp(-2b\ell)$.*

Proof. In view of Lemma 7.4, and using the definition of $\widehat{\mathfrak{D}}^{\ell}_{\boldsymbol{ce}}$, the inequalities in (7.4), the degree distributions in (7.8), and the analytic regularity (7.14), we conclude that $\Upsilon_{\widehat{\mathfrak{D}}^{\ell}_{\boldsymbol{ce}}}[\eta^{\|}] \lesssim \boldsymbol{p}^8_{\max}\big(S_1 + S_2\big)$, where the sums S_1 and S_2 are given by

$$S_1 = \sum_{j=2}^{\ell+1}\sum_{i=2}^{j} \Psi_{p^{\|}_j-1, s^{\|}_j-1}\sigma^{2(\ell+1-j)b_{\boldsymbol{c}}} C^{2s^{\|}_j}\Gamma(s^{\|}_j+3)^2,$$

$$S_2 = \sum_{j=2}^{j}\sum_{i=2}^{j} \Psi_{p^{\|}_j-1, s^{\|}_j-1}\sigma^{2(\ell+1-i)b_{\boldsymbol{e}}}\sigma^{2(\ell+1-j)(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})} C^{2s^{\|}_j}\Gamma(s^{\|}_j+4)^2.$$

The terms in the first sum S_1 are independent of the inner index i. Hence, by interpolation to real parameters $s^{\|}_j \in [3, p^{\|}_j]$ as in [11, Lemma 5.8], by applying [11, Lemma 5.12], and after possibly adjusting constants, we conclude $S_1 \lesssim \ell\exp(-2b_1(\ell+1)) \lesssim \exp(-2b_2\ell)$. The second sum S_2 can be estimated in exactly the same manner as the sum $S^{\|}$ in the proof of [11, Proposition 5.17], and we obtain $S_2 \lesssim \exp(-2b_3\ell)$. Adjusting the constants to absorb the algebraic factor $\boldsymbol{p}^8_{\max}$ yields the assertion. ☐

7.3.2. *Exponential convergence of $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},i}}$.* In this subsection, we bound the terms $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},i}}$ in (7.17), and first establish the following bounds for $\eta^{\perp}$, by using the properties (5.9) of the quasi-interpolation operator $\mathcal{I}^{\perp}_1$ for $\mathfrak{K} = K^{\perp}$.

Lemma 7.6. *Let $K = K^{\perp} \times K^{\|}_j$, $j \geq 2$, be an element in the terminal layer $\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$ of the form (7.7). For $s = 0, 1$, there holds*

$$(h^{\|}_K)^{2(s-1)}\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\eta^{\perp}\|^2_{L^2(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}}, b_{\boldsymbol{c}}\}\ell}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}, \qquad |\boldsymbol{\alpha}^{\perp}| = s, \tag{7.24}$$

and

$$(h^{\|}_K)^{2(s-1)}\|\mathsf{D}^{s}_{\|}\eta^{\perp}\|^2_{L^2(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}}, b_{\boldsymbol{c}}\}\ell}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\,. \tag{7.25}$$

Proof. To show (7.24), we apply (5.9), to get

$$(h^{\|}_K)^{2(s-1)}\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\eta^{\perp}\|^2_{L^2(K)} \lesssim (h^{\|}_K)^{2s-2}(h^{\perp}_K)^{4-2s-2(1-b_{\boldsymbol{e}})}\sum_{|\boldsymbol{\alpha}^{\perp}|=2}\|r^{1-b_{\boldsymbol{e}}}_{\boldsymbol{e}}\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}u\|^2_{L^2(K)}, \qquad |\boldsymbol{\alpha}^{\perp}| = s\,.$$

The application of the equivalences (7.4) implies that

$$\sum_{|\boldsymbol{\alpha}^{\perp}|=2}\|r^{1-b_{\boldsymbol{e}}}_{\boldsymbol{e}}\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}u\|^2_{L^2(K)} \lesssim (h^{\|}_K)^{-2(b_{\boldsymbol{e}}-b_{\boldsymbol{c}})}\|r^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}}_{\boldsymbol{c}}r^{1-b_{\boldsymbol{e}}}_{\boldsymbol{e}}\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}u\|^2_{L^2(K)} \lesssim (h^{\|}_K)^{-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}}}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\,.$$

Thus, combining these estimates and expressing the mesh sizes in terms of σ, cp. (7.5), (7.7), we see that, for $|\boldsymbol{\alpha}^{\perp}| = s$,

$$\begin{aligned}
&(h_K{}^{\|})^{2(s-1)}\|\mathsf{D}^{\boldsymbol{\alpha}^{\perp}}_{\perp}\eta^{\perp}\|^2_{L^2(K)} \\
&\quad\lesssim (h^{\|}_K)^{2s-2-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}}}(h^{\perp}_K)^{2-2s+2b_{\boldsymbol{e}}}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)} \simeq \sigma^{(\ell+1-j)(2s-2-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}})}\sigma^{\ell(2-2s+2b_{\boldsymbol{e}})}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)} \\
&\quad\simeq \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}\sigma^{2j(1-s)+2(s-1)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}|u|^2 \lesssim \sigma^{2\min\{b_{\boldsymbol{c}}, b_{\boldsymbol{c}}\}\ell}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\,.
\end{aligned}$$

To prove (7.25), we proceed similarly and obtain

$$\begin{aligned}
(h_K^{\parallel})^{2(s-1)}\|\mathsf{D}_{\parallel}^{s}\eta^{\perp}\|_{L^2(K)}^2 &\lesssim (h_K^{\parallel})^{2s-2}(h_K^{\perp})^{4-2(1-b_{\boldsymbol{e}})}\sum_{|\boldsymbol{\alpha}^{\perp}|=2}\|r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}\mathsf{D}_{\parallel}^{s}u\|_{L^2(K)}^2\\
&\lesssim (h_K^{\parallel})^{-2-2(b_{\boldsymbol{e}}-b_{\boldsymbol{c}})}(h_K^{\perp})^{2+2b_{\boldsymbol{e}}}\sum_{|\boldsymbol{\alpha}^{\perp}|=2}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}+s}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}_{\perp}^{\boldsymbol{\alpha}^{\perp}}\mathsf{D}_{\parallel}^{s}u\|_{L^2(K)}^2\\
&\lesssim \sigma^{(\ell+1-j)(-2-2b_{\boldsymbol{e}}+2b_{\boldsymbol{c}})}\sigma^{2\ell(1+b_{\boldsymbol{e}})}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\\
&\lesssim \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}\sigma^{2(j-1)}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\\
&\lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{N}^{s+2}_{-1-\boldsymbol{b}}(K)}\,.
\end{aligned}$$

This completes the proof. □

As a consequence of the preceding lemma, we have the following approximation bound in perpendicular direction.

Proposition 7.7. *Let $u\in\widehat{N}^4_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as defined in (7.12). Then there holds $\Upsilon_{\mathfrak{T}^{\ell}_{\boldsymbol{e},1}}[\eta^{\perp}]\leq C\exp(-2b\ell)$, for constants $C,b>0$ independent of ℓ.*

Proof. From Lemma 7.6 we find that, for $K\in\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$,

$$\begin{aligned}
(h_K^{\parallel})^{-2}\|\eta^{\perp}\|_{L^2(K)}^2+\|\mathsf{D}_{\perp}\eta^{\perp}\|_{L^2(K)}^2+\|\mathsf{D}_{\parallel}\eta^{\perp}\|_{L^2(K)}^2+(h_K^{\parallel})^2\|\mathsf{D}_{\parallel}^2\eta^{\perp}\|_{L^2(K)}^2&\\
\lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}\|u\|^2_{\widehat{N}^4_{-1-\boldsymbol{b}}(K)}&.
\end{aligned}$$

The assertion now follows by summing this estimate over all elements $K\in\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$, and by suitably adjusting constants. □

Moreover, for the approximation error $\eta^{\parallel}$ in parallel direction to edge $\boldsymbol{e}$, a similar estimate holds.

Proposition 7.8. *Let $u\in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as in (7.12), (7.14). Then, for ℓ sufficiently large, there holds $\Upsilon_{\mathfrak{T}^{\ell}_{\boldsymbol{e},1}}[\eta^{\parallel}]\leq C\exp(-2b\ell)$, for constants $b,C>0$ which are independent of $\ell\geq 1$.*

Proof. We note that, by (7.14), (7.13), the functions u and $\mathsf{D}_{\perp}u$ satisfy, respectively,

$$\begin{aligned}
\|r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+\alpha^{\parallel}}\mathsf{D}_{\parallel}^{\alpha^{\parallel}}u\|_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} &\leq C^{\alpha^{\parallel}+1}\Gamma(\alpha^{\parallel}+1), && \alpha^{\parallel}\geq 2,\\
\|r_{\boldsymbol{c}}^{-b_{\boldsymbol{c}}+\alpha^{\parallel}}\mathsf{D}_{\parallel}^{\alpha^{\parallel}}\mathsf{D}_{\perp}u\|_{L^2(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})} &\leq C^{\alpha^{\parallel}+2}\Gamma(\alpha^{\parallel}+2), && \alpha^{\parallel}\geq 2\,.
\end{aligned}$$

In view of (7.5), (7.7), these properties correspond to the one-dimensional analytic regularity assumptions (6.8) and (6.12), respectively. Moreover, due to (7.8), the polynomial degrees $p_K^{\parallel}$ are $\mathfrak{s}$-linearly increasing away from the corner $\boldsymbol{c}$. Hence, Proposition 6.2 respectively Proposition 6.3, and the tensor product structure of the elements yield

$$\sum_{K\in\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}}\Big((h_K^{\parallel})^{-2}\|\eta^{\parallel}\|_{L^2(K)}^2+\|\mathsf{D}_{\parallel}\eta^{\parallel}\|_{L^2(K)}^2+(h_K^{\parallel})^2\|\mathsf{D}_{\parallel}^2\eta^{\parallel}\|_{L^2(K)}^2\Big)\lesssim\exp(-2b\ell),$$

respectively, $\sum_{K\in\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}}\|\mathsf{D}_{\perp}\eta^{\parallel}\|_{L^2(K)}^2\lesssim\exp(-2b\ell)$. This completes the proof. □

Finally, we bound the term in $\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},2}}[\eta]$.

Proposition 7.9. *Let u be in $\widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^{\ell}_{\boldsymbol{ce}})$ as defined in (7.12).*

(1) *For $K\in\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e}}$, there holds:*

$$T^K_{\boldsymbol{e},2}[\eta]\lesssim(h_K^{\perp})^{2b_{\boldsymbol{e}}}(h_K^{\parallel})^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}_{\perp}^2u\|_{L^2(K)}^2\,. \tag{7.26}$$

(2) *Moreover,*

$$\Upsilon_{\widehat{\mathfrak{T}}^{\ell}_{\boldsymbol{e},2}}[\eta]\leq C\exp(-2b\ell), \tag{7.27}$$

with constants $b,C>0$ independent of ℓ.

Proof. To show (7.26), we note that, by Hölder's inequality and due to the fact that $b_{\boldsymbol{c}}, b_{\boldsymbol{e}} \in (0,1)$,

$$\sum_{|\boldsymbol{\alpha}^\perp|=2} \|\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta\|^2_{L^1(K)} \lesssim \|r_{\boldsymbol{c}}^{-1+b_{\boldsymbol{c}}}\rho_{\boldsymbol{ce}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \sum_{|\boldsymbol{\alpha}^\perp|=2} \|r_{\boldsymbol{c}}^{1-b_{\boldsymbol{c}}}\rho_{ce}^{\max\{1-b_{\boldsymbol{e}},0\}}\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta\|^2_{L^2(K)}$$
$$\leq \|r_{\boldsymbol{c}}^{b_{\boldsymbol{c}}-b_{\boldsymbol{e}}}r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \sum_{|\boldsymbol{\alpha}^\perp|=2} \|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta\|^2_{L^2(K)}\,.$$

Then, employing (7.5) in direction parallel to $\boldsymbol{e}$ yields that we have $\|r_{\boldsymbol{c}}^{b_{\boldsymbol{c}}-b_{\boldsymbol{e}}}r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \simeq (h_K^\parallel)^{2b_{\boldsymbol{c}}-2b_{\boldsymbol{e}}}\|r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)}$. Since $|K| \simeq h_K^\parallel (h_K^\perp)^2$, we further have $\|r_{\boldsymbol{e}}^{-1+b_{\boldsymbol{e}}}\|^2_{L^2(K)} \lesssim |K|(h_K^\perp)^{2b_{\boldsymbol{e}}-2}$. Furthermore, for $|\boldsymbol{\alpha}^\perp| = 2$, noting that $\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\eta = \mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}u - \Pi^\parallel_{p_K^\parallel}(\mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}\mathcal{I}_1^\perp u) = \mathsf{D}_\perp^{\boldsymbol{\alpha}^\perp}u$ (since $\mathcal{I}_1^\perp u \in \mathbb{P}_1(K^\perp)$) implies (7.26).

To prove the bound (7.27), we refer to (7.26), (7.5), and (7.7). This results in

$$T^K_{\boldsymbol{e},2}[\eta] \lesssim \sigma^{2\ell b_{\boldsymbol{e}}}\sigma^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})(\ell+1-j)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)} = \sigma^{2b_{\boldsymbol{c}}(\ell+1-j)+2b_{\boldsymbol{e}}(j-1)}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}$$
$$\lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{N}^2_{-1-\boldsymbol{b}}(K)}\,.$$

Summing this last bound over all elements $K \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ yields the assertion. □

7.3.3. *Conclusion of proof of* (7.17). The proof of the exponential convergence bound (7.17) on the hp-dG interpolation error η on $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}$ in Theorem 7.1 follows now straightforwardly by estimating the terms on the left-hand side of (7.17) using the above results.

The proof of exponential convergence (7.17) on $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}$ and on $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{c}}$ claimed in Theorem 7.1 follows from the bound in the corner-edge patch $\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}$ upon noticing (7.10), (7.11).

7.4. **Exponential convergence estimates in elements at Dirichlet edges.** Next, we consider the case where $\boldsymbol{e} \in \mathcal{E}_D$ is a Dirichlet edge, i.e., $e \in \mathcal{E}_D$, and establish the analog of Theorem 7.1. According to (2.11) and [3], the solution regularity is characterized by the *homogeneous corner-edge seminorms*

$$|u|^2_{\widehat{M}^{|\boldsymbol{\alpha}|}_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|}\rho_{\boldsymbol{ce}}^{-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|}\mathsf{D}^{\boldsymbol{\alpha}}u\right\|^2_{L^2(\widehat{\Omega}^\ell_{\boldsymbol{ce}})}, \qquad k > k_{\boldsymbol{\beta}}. \tag{7.28}$$

While exponential convergence for solutions with regularity in this family of spaces was already shown in [11], we present an alternative argument, based on the preceding analysis of the Neumann case. We say a function $u \in H^1(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ belongs to $A_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ if $u \in \widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$, for $k > k_{\boldsymbol{\beta}}$, and there is a constant $d_u > 0$ such that

$$\|u\|_{\widehat{M}^k_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} \leq d_u^{k+1}k!\,, \qquad \forall k > k_{\boldsymbol{\beta}}\,. \tag{7.29}$$

Corollary 7.10. *Let $\boldsymbol{e} \in \mathcal{E}_D$ be a Dirichlet edge. Let $u \in A_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ as in* (7.28), (7.29)*, and let Π_K be the elemental approximation operators chosen in accordance with Section 5.1. Then for η, $\eta^\perp$, $\eta^\parallel$ as in* (5.15), (5.16)*, and for ℓ sufficiently large, there holds*

$$\Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}[\eta^\perp] + \Upsilon_{\widehat{\mathfrak{O}}^\ell_{\boldsymbol{ce}}}[\eta^\parallel] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\perp] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},1}}[\eta^\parallel] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},2}}[\eta] \leq C\exp(-2b\ell), \tag{7.30}$$

with constants $b, C > 0$ independent of ℓ.

In addition, there holds

$$\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},D}}[\eta^\perp] + \Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e},D}}[\eta^\parallel] \leq C\exp(-2b\ell), \tag{7.31}$$

with constants $b, C > 0$ independent of ℓ.

Proof. For every $k \geq 0$, there holds $|u|_{\widehat{N}^k(\widehat{\Omega}^\ell_{\boldsymbol{ce}})} \leq |u|_{\widehat{M}^k(\widehat{\Omega}^\ell_{\boldsymbol{ce}})}$. Hence, $u \in A_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$ implies $u \in B_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$, and the bound (7.30) follows from Theorem 7.1.

To bound (7.31), let K be in $\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$. Then, by (5.9), the definition of the corner-edge seminorm (7.28), and the properties (7.5), we find that

$$(h_K^\perp)^{-2}\|\eta^\perp\|^2_{L^2(K)} \lesssim (h_K^\perp)^{2b_{\boldsymbol{e}}}(h_K^\parallel)^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})}\|r_{\boldsymbol{c}}^{b_{\boldsymbol{e}}-b_{\boldsymbol{c}}}r_{\boldsymbol{e}}^{1-b_{\boldsymbol{e}}}\mathsf{D}^2_\perp u\|^2_{L^2(K)}$$
$$\lesssim (h_K^\perp)^{2b_{\boldsymbol{e}}}(h_K^\parallel)^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})}|u|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)}.$$

In direction parallel to edge $\boldsymbol{e}$, we proceed similarly: The stability of the L^2-projection and equations (7.28), (7.5), yield

$$(h_K^\perp)^{-2}\|\eta^\parallel\|^2_{L^2(K)} \lesssim (h_K^\perp)^{-2}\|u\|^2_{L^2(K)} \lesssim (h_K^\perp)^{2b_{\boldsymbol{e}}}(h_K^\parallel)^{2(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})}|u|^2_{\widehat{M}^0_{-1-\boldsymbol{b}}(K)}.$$

Therefore, expressing the mesh sizes in term of σ, cp. (7.5), implies

$$(h_K^\perp)^{-2}\big(\|\eta^\perp\|^2_{L^2(K)} + \|\eta^\parallel\|^2_{L^2(K)}\big) \lesssim \sigma^{2b_{\boldsymbol{e}}\ell}\sigma^{2(\ell+1-j)(b_{\boldsymbol{c}}-b_{\boldsymbol{e}})}\|u\|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)}$$
$$\lesssim \sigma^{2(\ell+1-j)b_{\boldsymbol{c}}+2(j-1)b_{\boldsymbol{e}}}|u|^2_{\widehat{M}^0_{-1-\boldsymbol{b}}(K)} \lesssim \sigma^{2\min\{b_{\boldsymbol{c}},b_{\boldsymbol{e}}\}\ell}|u|^2_{\widehat{M}^2_{-1-\boldsymbol{b}}(K)}\,.$$

Summing the above bound over all elements in $\widehat{\mathfrak{T}}^\ell_{\boldsymbol{e}}$ implies the asserted exponential convergence bound. $\square$

7.5. **Exponential convergence at corner elements.** To conclude the proof of Theorem 5.6 it remains to show exponential convergence in elements $K_{\boldsymbol{c}} \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}$ which abut at a corner $\boldsymbol{c} \in \mathcal{C}$ of Ω so that $\overline{K}_{\boldsymbol{c}} \cap \boldsymbol{c} \neq \emptyset$. Such elements $K_{\boldsymbol{c}}$ are shape-regular and axiparallel, with diameter $h_{\boldsymbol{c}} = \mathcal{O}(\sigma^\ell)$. We are left to bound the term $T^{K_{\boldsymbol{c}}}_{\boldsymbol{c}}[\eta]$ defined in (5.25). On $K_{\boldsymbol{c}}$, we use the quasi-interpolant $\mathcal{I}_1$ defined in (5.6) for $\mathfrak{K} = K_{\boldsymbol{c}}$. Then

$$\eta|_{K_{\boldsymbol{c}}} = u|_{K_{\boldsymbol{c}}} - \mathcal{I}_1(u|_{K_{\boldsymbol{c}}})\,. \tag{7.32}$$

The quasi-interpolant $\mathcal{I}_1$ is well-defined under the (minimal) regularity $u \in W^{1,1}(K_{\boldsymbol{c}})$. Furthermore, by (5.8) there holds that $\|\eta\|_{L^2(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}\|\nabla\eta\|_{L^2(K_{\boldsymbol{c}})}$, and $\|\nabla\eta\|_{L^2(K_{\boldsymbol{c}})} = \|\nabla u - \boldsymbol{\Pi}_0\nabla u\|_{L^2(K_{\boldsymbol{c}})}$. We conclude

$$T^{K_{\boldsymbol{c}}}_{\boldsymbol{c}}[\eta] \lesssim \|\nabla u - \boldsymbol{\Pi}_0\nabla u\|^2_{L^2(K_{\boldsymbol{c}})} + h_{\boldsymbol{c}}^{-1}|u|^2_{W^{2,1}(K_{\boldsymbol{c}})}\,. \tag{7.33}$$

To bound the first term, applying standard approximations properties for Π_0 would imply a bound of order $\mathcal{O}(h_{\boldsymbol{c}})$ *provided* $u \in H^2(K_{\boldsymbol{c}})$. The weaker regularity $u \in N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\{\boldsymbol{c}\},\emptyset)$ suffices to obtain a (slightly weaker, yet still exponentially convergent) bound, due to the embedding $N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\{\boldsymbol{c}\},\emptyset) \subset H^1(K_{\boldsymbol{c}})$ being compact. The next two lemmas provide an exponential bound on the first term in (7.33).

Lemma 7.11. *For corner weight parameters $b_{\boldsymbol{c}} \in (0, 1/2)$, and edge weight parameters $b_{\boldsymbol{e}} \in (0,1)$, for $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}} \subset \mathcal{E}_N$, $\boldsymbol{c} \in \mathcal{C}$, we have the compact embeddings*

$$N^1_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\{\boldsymbol{c}\},\emptyset) \subset L^2(K_{\boldsymbol{c}}), \qquad N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\{\boldsymbol{c}\},\emptyset) \subset H^1(K_{\boldsymbol{c}})\,. \tag{7.34}$$

Proof. We note that, in the $N_{\boldsymbol{\beta}}$-spaces above, all edges $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ are Neumann edges (although all that follows will hold verbatim if only some $\boldsymbol{e} \in \mathcal{E}_{\boldsymbol{c}}$ belong to $\mathcal{E}_N$). We write, for simplicity, $N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}})$ in place of $N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\{\boldsymbol{c}\},\emptyset)$. The key observation of the proof is the equivalence $N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}}) \simeq \boldsymbol{H}^{2,2}_{\beta_m,\beta_{ij}}(K_{\boldsymbol{c}})$ proved in [8, Section 2] for the indicated range of weight exponents $\beta_m \in (0, 1/2)$ and $\beta_{ij} \in (0,1)$ (cp. (2.10) and Remark 2.2).

Then, [8, Theorem 3.8] implies that $H^{1+\theta}(K_{\boldsymbol{c}}) \supset \boldsymbol{H}^{2,2}_{\beta_m,\beta_{ij}}(K_{\boldsymbol{c}}) \simeq N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}})$, with *continuous* embedding, provided that $\theta := 1 - \max\{\beta_m, \beta_{ij}\} > \varepsilon$, for sufficiently small $\varepsilon > 0$. Using that $\beta_{ij} = \beta_{\boldsymbol{e}} + 2$ and $\beta_m = \beta_{\boldsymbol{c}} + 2$ (cp. Remark 2.2), we obtain $\theta \in (0,1)$ if and only if $0 < b_{\boldsymbol{e}} < 1$ and $0 < b_{\boldsymbol{c}} < 1/2$, cp. (2.10), which is the asserted range of corner and edge weight exponents. The compactness of the second embedding in (7.34) now follows from the fact that it is a composition of the continuous embedding $N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}}) \subset H^{1+\theta}(K_{\boldsymbol{c}})$ and the compact (by Rellich's Theorem) injection $H^{1+\theta}(K_{\boldsymbol{c}}) \subset H^1(K_{\boldsymbol{c}})$ for $\theta > 0$. The compactness of the first embedding in (7.34) follows analogously. $\square$

Lemma 7.12. *Let $u \in N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)$, with edge weights $\beta_{\boldsymbol{e}} \in (-2,-1)$ and corner weights $\beta_{\boldsymbol{c}} \in (-3/2,-1)$. Then for the quasi-interpolant $\mathcal{I}_1$ for $\mathfrak{K} = K_{\boldsymbol{c}}$ in (5.6), there exists $C>0$ independent of $h_{\boldsymbol{c}} \in (0,1]$ and of u such that $\|\nabla(u-\mathcal{I}_1 u)\|_{L^2(K_{\boldsymbol{c}})} = \|\nabla u - \boldsymbol{\Pi}_0 \nabla u\|_{L^2(K_{\boldsymbol{c}})} \le C h_{\boldsymbol{c}}^{b_{\boldsymbol{c}}} |u|_{N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)}$.*

Proof. We observe that $u \in N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)$ implies that $\nabla u \in N^1_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)^3$. We denote $\boldsymbol{v} = \nabla u \in N^1_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)^3$. Observe that $\boldsymbol{\Pi}_0(\boldsymbol{v})$ is the (componentwise) average of $\boldsymbol{v}$ over $K_{\boldsymbol{c}}$. From the compactness of the embedding $N^1_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)^3 \subset L^2(K_{\boldsymbol{c}})^3$ in Lemma 7.11, we proceed along the lines of [13, Section A.2.4] and use appropriate scaling (in particular, recalling that $-2 < \beta_e < -1$ implies that, for $k=1$ in the sixth term of (2.11), the inhomogeneous weight exponents $\beta_{\boldsymbol{e}} + |\boldsymbol{\alpha}^\perp| < 0$) to conclude that there exists a constant $C>0$ independent of $\boldsymbol{v}$ such that $\|\boldsymbol{v} - \boldsymbol{\Pi}_0(\boldsymbol{v})\|_{L^2(K_{\boldsymbol{c}})} \le C h_{\boldsymbol{c}}^{-1-\beta_{\boldsymbol{c}}} |\boldsymbol{v}|_{N^1_{\boldsymbol{\beta}}(K_{\boldsymbol{c}})}$. Referring to (2.10) completes the proof. □

It remains to bound the term $h_{\boldsymbol{c}}^{-1}|u|^2_{W^{2,1}(K_{\boldsymbol{c}})}$ in (7.33).

Lemma 7.13. *Let $u \in N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)$, with some $\beta_{\boldsymbol{e}} \in (-2,-1)$, and with some $\beta_{\boldsymbol{c}} \in (-3/2,-1)$, and for $K_{\boldsymbol{c}} \in \widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}$ with $\overline{K}_{\boldsymbol{c}} \cap \boldsymbol{c} \neq \emptyset$, and $\overline{K}_{\boldsymbol{c}} \cap \boldsymbol{e} \neq \emptyset$ for $\boldsymbol{c} \subset \overline{\boldsymbol{e}}$, for some $\boldsymbol{e} \in \mathcal{E}_N$. Then, for any $0 < h_{\boldsymbol{c}} = \mathrm{diam}(K_{\boldsymbol{c}}) \le 1$, there holds*

$$|u|_{W^{2,1}(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}} |u|_{N^2_{\boldsymbol{\beta}}(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)}\,. \tag{7.35}$$

Here, $b_{\boldsymbol{c}} = -1-\beta_{\boldsymbol{c}} \in (0,1/2)$ is as in (2.10).

Proof. We may assume that $K_{\boldsymbol{c}} \cap \omega_{\boldsymbol{e}} = \emptyset$. There holds

$$\begin{aligned}
|u|_{W^{2,1}(K_{\boldsymbol{c}})} &= \sum_{|\boldsymbol{\alpha}|=2} \|\mathsf{D}^{\boldsymbol{\alpha}}\eta\|_{L^1(K_{\boldsymbol{c}})} = \sum_{|\boldsymbol{\alpha}|=2} \|\mathsf{D}^{\boldsymbol{\alpha}}\eta\|_{L^1(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{c}})} + \sum_{|\boldsymbol{\alpha}|=2} \|\mathsf{D}^{\boldsymbol{\alpha}}\eta\|_{L^1(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{ce}})} \\
&\le \sum_{|\boldsymbol{\alpha}|=2} \left\| r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{c}})} \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \mathsf{D}^{\boldsymbol{\alpha}}\eta \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{c}})} \\
&\quad + \sum_{|\boldsymbol{\alpha}|=2} \left\| r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{-\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0)} \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{ce}})} \\
&\qquad \times \left\| r_{\boldsymbol{c}}^{-1-b_{\boldsymbol{c}}+|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0)} \mathsf{D}^{\boldsymbol{\alpha}}\eta \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{ce}})}.
\end{aligned}$$

We note that, for $0 \le |\boldsymbol{\alpha}| \le 2$, there holds $\left\| r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{5/2+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}$, and similarly, $\left\| r_{\boldsymbol{c}}^{1+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \rho_{\boldsymbol{ce}}^{-\max(-1-b_{\boldsymbol{e}}+|\boldsymbol{\alpha}^\perp|,0)} \right\|_{L^2(K_{\boldsymbol{c}}\cap\omega_{\boldsymbol{ce}})} \lesssim h_{\boldsymbol{c}}^{5/2+b_{\boldsymbol{c}}-|\boldsymbol{\alpha}|} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}}$. We arrive at $\sum_{|\boldsymbol{\alpha}|=2} \|\mathsf{D}^{\boldsymbol{\alpha}}\eta\|_{L^1(K_{\boldsymbol{c}})} \lesssim h_{\boldsymbol{c}}^{1/2+b_{\boldsymbol{c}}} |u|_{N^2(K_{\boldsymbol{c}};\mathcal{C},\mathcal{E}_D)}$ which completes the proof. □

Inserting the estimates in the previous lemmas into (7.33), we arrive at the following exponential convergence result in corner elements.

Proposition 7.14. *Let $u \in \widehat{N}^2_{-1-\boldsymbol{b}}(\widehat{\Omega}^\ell_{\boldsymbol{ce}})$, with $\boldsymbol{b}$ as in Remark 2.4. Then, there exist constants $b, C > 0$ such that $\Upsilon_{\widehat{\mathfrak{T}}^\ell_{\boldsymbol{c}}}[\eta] \le C \exp(-2b\ell)|u|^2_{N^2(\Omega;\mathcal{C},\mathcal{E}_D)}$.*

7.6. **Proof of Theorem 5.6.** The exponential convergence of *hp*-dGFEM, Theorem 5.6, follows now immediately from the quasi-optimality results, Theorem 5.5, and from the fact that, by our analysis in Section 7, all terms on the right-hand side of the estimate (5.28) convergence exponentially with respect to the number of mesh layers ℓ. Furthermore, for the number of degrees of freedom in either of the *hp*-dG spaces in (3.10) and (3.11) there holds $N \simeq \ell^5 + \mathcal{O}(\ell^4)$, which yields the desired estimate (5.32).

Remark 7.15. We note that Theorem 5.6 remains true in the pure Neumann case. Indeed, the *hp*-approximation analysis on geometric meshes presented in this work as applied to the *hp*-dGFEM (4.2) with $\mathcal{F}_D(\mathcal{M}) = \emptyset$ and based on the *hp*-space $V(\mathcal{M},\boldsymbol{\Phi},\boldsymbol{p})/\mathbb{R}$ leads to the bound (5.32) as well. This simply follows from the fact that all the interpolants in our error analysis reproduce constant functions.

References

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2001.

[2] I. Babuška and B. Q. Guo. Regularity of the solution of elliptic problems with piecewise analytic data. part I. Boundary value problems for linear elliptic equation of second order. *SIAM J. Math. Anal.*, 19:172–203, 1988.

[3] M. Costabel, M. Dauge, and S. Nicaise. Analytic regularity for linear elliptic systems in polygons and polyhedra. *Math. Models Methods Appl. Sci.*, 22(8), 2012.

[4] M. Costabel, M. Dauge, and Ch. Schwab. Exponential convergence of *hp*-FEM for Maxwell's equations with weighted regularization in polygonal domains. *Math. Models Methods Appl. Sci.*, 15(4):575–622, 2005.

[5] V. Dolejší, M. Feistauer, and Ch. Schwab. A finite volume discontinuous Galerkin scheme for nonlinear convection-diffusion problems. *Calcolo*, 39(1):1–40, 2002.

[6] W. Gui and I. Babuška. The *h*, *p* and *h-p* versions of the finite element method in 1 dimension. II. The error analysis of the *h*- and *h-p* versions. *Numer. Math.*, 49(6):613–657, 1986.

[7] B. Q. Guo. The *h-p* version of the finite element method for solving boundary value problems in polyhedral domains. In *Boundary Value Problems and Integral Equations in Nonsmooth Domains*, volume 167 of *Lecture Notes in Pure and Appl. Math.*, pages 101–120. Dekker, New York, 1995.

[8] B. Q. Guo and I. Babuška. Regularity of the solutions for elliptic problems on nonsmooth domains in $\mathbb{R}^3$. I. Countably normed spaces on polyhedral domains. *Proc. Roy. Soc. Edinburgh Sect. A*, 127(1):77–126, 1997.

[9] D. Schötzau and Ch. Schwab. Exponential convergence in a Galerkin least squares *hp*-FEM for Stokes flow. *IMA J. Numer. Anal.*, 21:53–80, 2001.

[10] D. Schötzau, Ch. Schwab, and T. P. Wihler. *hp*-dGFEM for Second Order Elliptic Problems in Polyhedra I: Stability on Geometric Meshes. *SIAM J. Numer. Anal.*, 51(3):1610–1633, 2013.

[11] D. Schötzau, Ch. Schwab, and T. P. Wihler. *hp*-dGFEM for Second Order Elliptic Problems in Polyhedra II: Exponential Convergence. *SIAM J. Numer. Anal.*, 51(4):2005–2035, 2013.

[12] Ch. Schwab. *p- and hp-FEM – Theory and Application to Solid and Fluid Mechanics*. Oxford University Press, Oxford, 1998.

[13] T. P. Wihler. *Discontinuous Galerkin FEM for Elliptic Problems in Polygonal Domains*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2002. Diss. ETH No. 14973.

[14] T. P. Wihler and M. Wirz. Mixed *hp*-discontinuous Galerkin FEM for linear elasticity and Stokes flow in three dimensions. *Math. Models Methods Appl. Sci.*, 22(8), 2012.

Mathematics Department, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: `schoetzau@math.ubc.ca`

Seminar of Applied Mathematics, ETH Zürich, 8092 Zürich, Switzerland
E-mail address: `schwab@math.ethz.ch`

Mathematisches Institut, Universität Bern, 3012 Bern, Switzerland
E-mail address: `wihler@math.unibe.ch`